Paper 3740-2015

Risk-Adjusting Provider Performance Utilization Metrics

Tracy B. Lewis, BCBSNC; Stephanie Poley

ABSTRACT

Pay-for-performance programs are putting increasing pressure on providers to better manage patient utilization through care coordination, with the philosophy that good preventive services and routine care can prevent the need for some high-resource services. Evaluation of provider performance frequently includes measures such as acute care events (ER and inpatient), imaging, and specialist services, yet rarely are these indicators adjusted for the underlying risk of providers' patient panel. In part, this is because standard patient risk scores are designed to predict costs, not the probability of specific service utilization. As such, Blue Cross and Blue Shield of North Carolina has developed a methodology to model our members' risk of these events in an effort to ensure that providers are evaluated fairly and to prevent our providers from adverse selection practices. Our risk modeling takes into consideration members' underlying health conditions and limited demographic factors during the previous 12 month period, and employs two-part regression models using SAS® software. These risk-adjusted measures will subsequently be the basis of performance evaluation of primary care providers for our Accountable Care Organizations (ACOs) and medical home initiatives.

INTRODUCTION

Provider performance evaluations rely on utilization metrics to differentiate between providers on metrics sensitive to PCP care coordination. Historically, provider performance evaluation and profiling is not risk adjusted. Providers are compared to peers without taking into consideration some of the differences in their patients' health conditions. Differences in underlying risk can lead to inaccurate conclusions if metrics are not risk adjusted. Without proper risk adjustment, providers could be penalized for utilization based on a member's underlying risk. Risk adjustment levels the playing field and allows the variation to reflect the ways the members use services and how providers steer members to utilize services.

This paper outlines the use of a two-part regression model to create a member level risk score for ER use. It is important to note that for this model, an ER visit occurs when a member visits the ER but is not admitted during that visit.

METHODOLOGY AND RESULTS

The model takes into consideration members' underlying health conditions and limited demographic factors. The model is a retrospective regression model in that predictors in the model will be from the same 12 month timeframe as the ER utilization. This is different than a typical predictive model where you are predicting something in the future and often use past utilization as predictors. The ER risk score allows for risk adjustment to compare utilization across populations, provider groups and ACOs.

The first step in building the model is to determine the type of distribution for the data. Figure 1 shows the criteria for a zero inflated distribution as well as the difference between a Poisson and Negative Binomial distribution.

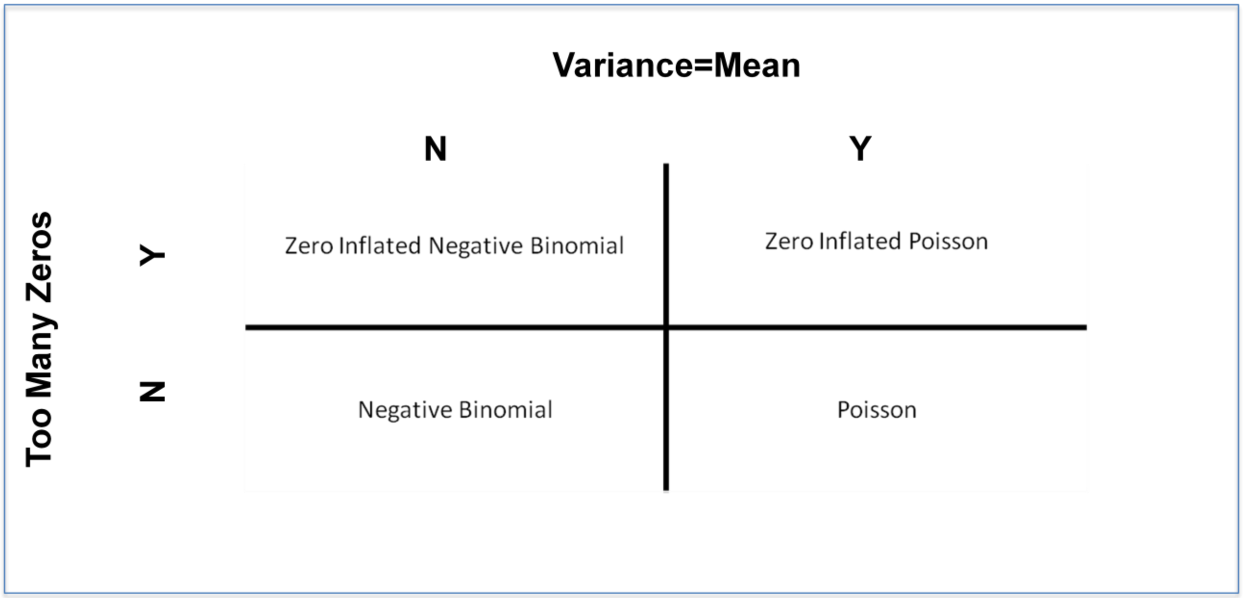

Figure 1. Count Model Distributions

There are two questions that need to be answered when looking at your data; are there too many zeros and does the variance equal the mean? Figure 2 shows the mean ER visit count is 0.15 visits per member per year and the variance is 0.30. Since the mean does not equal the variance and almost 90% of the members have no ER visits, the appropriate modeling distribution for ER visits is a zero inflated negative binomial.

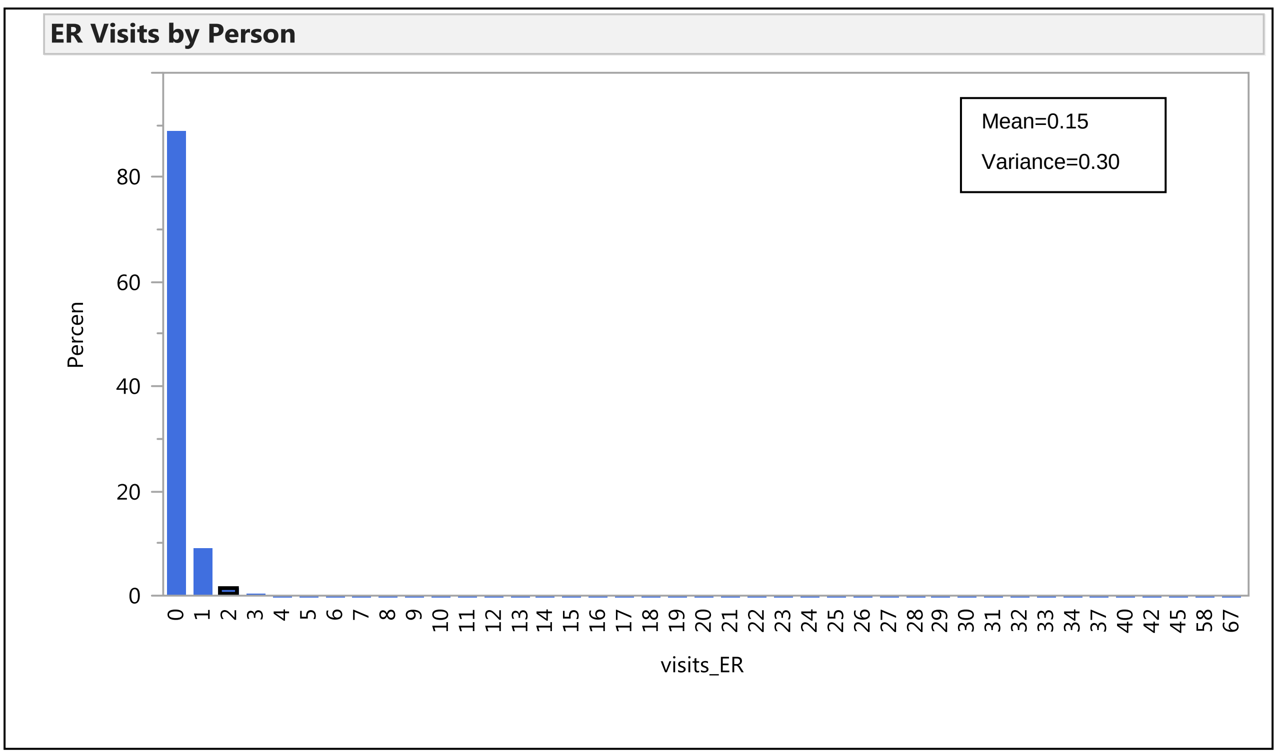

Figure 2. Distribution of ER Visits by Person

The model is built as a two-part regression model. Logistic regression predicts the probability of at least 1 ER visit. Negative Binomial model accounts for the second part of the model, the number of ER visits.

STEP ONE: VARIABLE SELECTION AND VARIABLE REDUCTION

When deciding on the model inputs or independent variables, you must decide on the variables to include. The purpose of the model is to adjust for underlying member health conditions so the list of potential input variables is limited to member level data.

Underlying member health conditions are defined using the risk markers in Symmetry ERG® (Episode Risk Groups®) and Symmetry PRG® (Pharmacy Risk Groups®) v7.6 products. Symmetry ERG uses medical and pharmacy claims to calculate member risk scores that predict total member cost for a specific 12 month period. Symmetry PRG uses pharmacy claims only to predict total member cost. It is important to note that these risk scores cannot be used to adjust utilization since they are built to predict cost. However, the risk markers used to compute these total cost risk scores are used in this model development. Symmetry ETG® (Episode Treatment Groups®) episodes are the building blocks for the 189 ERG risk markers. Each ERG risk marker is comprised of similar conditions and severity. There can be several risk markers with the same name but different levels assigned. For example, there are five levels of orthopedic trauma, fracture or dislocation, I to V. Each level groups ETG episodes of similar clinical and member risk and generally speaking the risk increases with each level. The risk markers do not take into account treatment, with the exception of oncology active treatment. For example, a patient with one physician visit to manage his migraines, a patient with 3 visits for migraine conditions and a patient on migraine prescriptions with no physician encounters all have the same ERG risk marker. Treatment is not included as you only want to take into account member underlying health conditions so as to not reward or penalize a provider for treatment decisions. The 153 PRG risk markers use NDC to DCC (Drug Class Code) mappings which in turn map to PRGs.

The other variables that are included are whether the member lives in a rural area as well as census variables including race, education and income. The introduction of these variables is to proxy for health education or health literacy. A member's health literacy is indicative of the understanding of his health and how it impacts the perception and need for particular services. The final variables we consider are age and gender.

Once you have the list of independent variables, you must apply data reduction techniques to reduce the number of inputs. The first step is to use PROC FACTOR to identify the factor structure underlying the data. The methodology developed by the ERG risk markers ensures that there is not more than one ERG in any given factor so no further data reduction technique was necessary for factors with a single ERG marker. For factors with only PRG markers, you can create a new factor variable to indicate if a member has any of the PRG markers. For factors with at least one PRG marker you might see other PRG markers and/or ERG markers grouping to the same factor. When there is a mix of ERG and PRG markers in the same factor, the ERG marker is used. For this model, after data reduction techniques were applied, there were 239 factors created from over 350 potential model variables.

STEP TWO: LOGISTIC MODEL (ER VISIT)

After applying variable data reduction technique, a forward stepwise logistic regression using PROC LOGISTIC with SELECTION=FORWARD is employed to select variables: Figure 3 demonstrates that the c-statistic levels out at ~0.80 and that the fit is maximized.

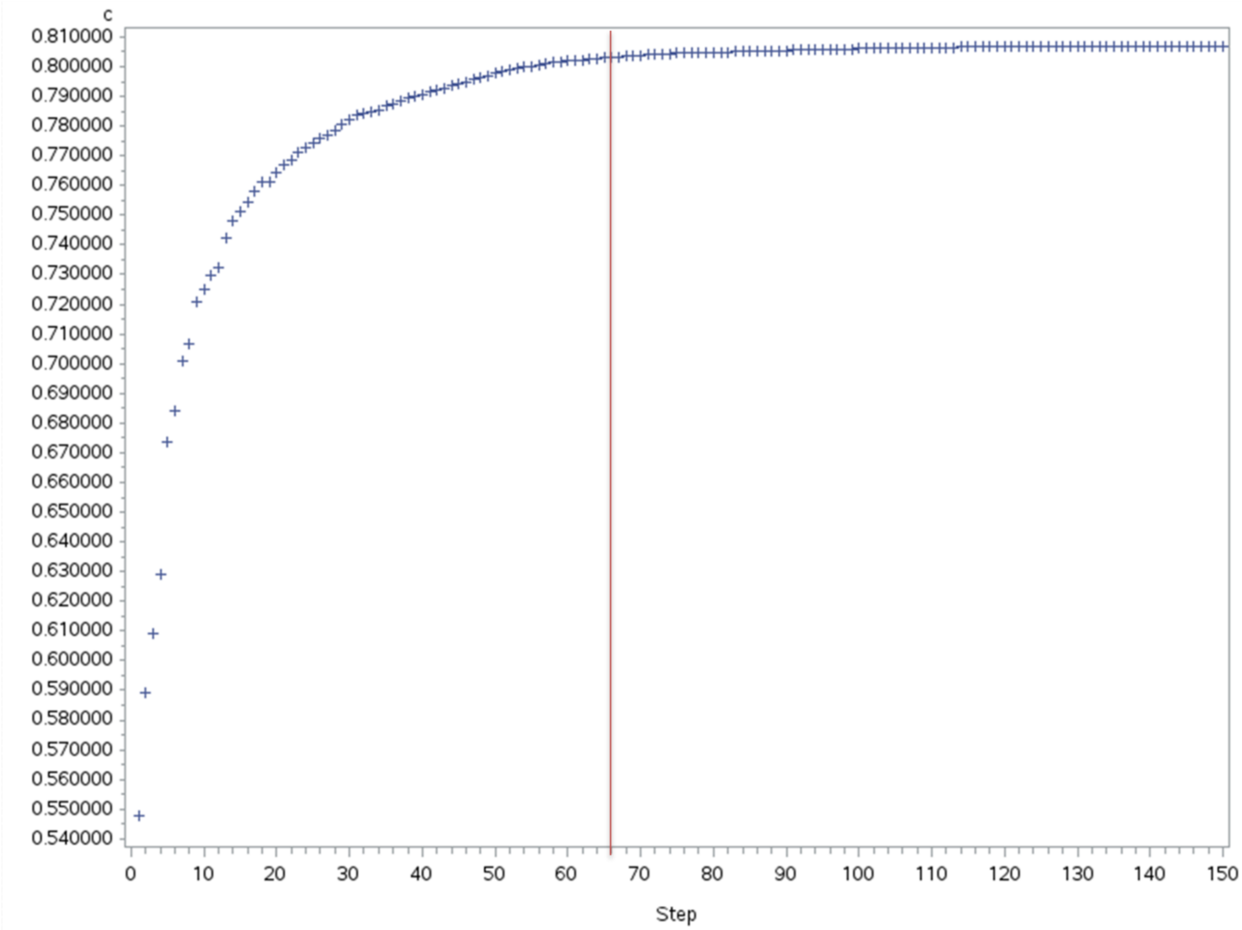

Figure 3. Forward Stepwise Logistic Regression C-statistic

A regression model, PROG REG, was then run on the variables selected from the stepwise regression to assess potential multicollinearity. The VIF (Variance Inflation Factor) was < 1.40 for all variables, which is below the standard of < 2.0 for regression models. Finally, PROG LOGISTIC is run on these variables to compute the probability of an ER visit, P1. Table 1 lists the top 10 predictors for the Logistic model.

Variable
Orthopedic trauma, fracture or dislocation, II
Orthopedic trauma, fracture or dislocation, III
Anti-emetic agents, not elsewhere classified, II OR Anti-emetic agents with noninsulin diabetic agents, selected steroids OR Anti-emetic agents, with CAD
Lower cost substance abuse
Lower cost cardiology, I
Member age
Low cost dermatology, I
Orthopedic trauma, fracture or dislocation, I
Lower cost gastroenterology, I
Lower cost neurology

Table 1. Logistic Top 10 Predictors

The top predictors are chosen based on the Wald Chi-Square statistic. Variables with larger Wald Chi-Square values have less variation so you can be confident these are strong predictors in the model.

STEP THREE: NEGATIVE BINOMIAL COUNT MODEL (NUMBER OF ER VISITS)

Once the variables for the logistic model have been determined, the next step is determining the variables for the count portion of the model. Originally, we planned to use PROC GENMOD with DIST=ZINB for our model. Unfortunately, we encountered multiple processing issues executing this procedure in our environment and were unable to successfully run this procedure. To overcome this obstacle, our team decided to run the two parts as separate models and combine the results. The logistic portion as outlined in step two is the first model and the second portion is the count model. In order to take into account the abundance of zeros in the count model, we developed the model on those members with at least 1 ER visit but ran the model on (ER visit count – 1). Figure 4 shows a side by side comparison of the ER distribution of the original dataset as well as the modified dataset. Both graphs show the excess of zero visits.

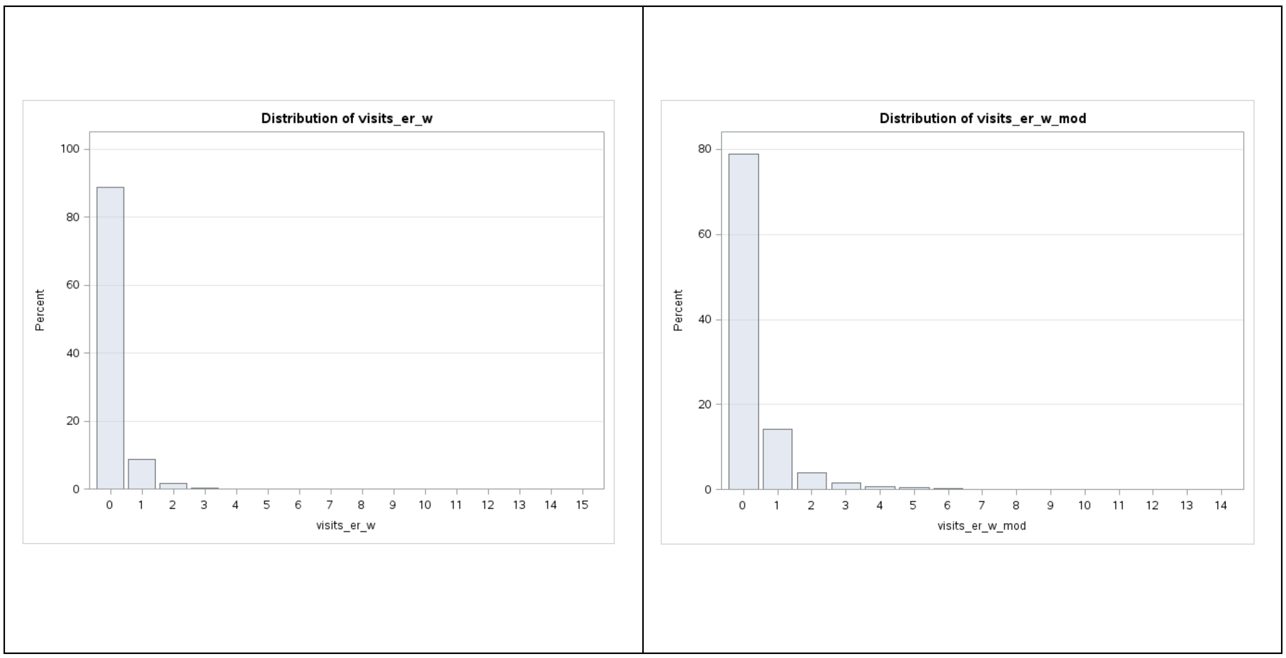

Figure 4. ER Visit Distribution

A few variable modifications were necessary for the negative binomial count model. Age/gender category variables were used instead of the member age and female flags used in the logistic model. In addition, all census variables were changed to have a 0 to 1 scale. Multiple iterations of the count model (PROC GENMOD with DIST=NEGBIN) were run, while comparing the AIC, to find the optimal variables to include in the model. Table 2 shows the top 10 predictors for the negative binomial count model.

Variables
Orthopedic trauma, fracture or dislocation, II
19 to 34 year old females
Lower cost substance abuse
19 to 34 year old males
Pct white - census variable
Pct high school - census variable
Lower cost dermatology, II
Asthma, COPD, II
35 to 44 year old females
Hypertension, with significant complication/comorbidity
Hypertension, without complication/comorbidity

Table 2. Negative Binomial Top 10 Predictors

Top Predictor Comparison

Table 3 compares the top 10 predictors for the Logistic and Count Models. Overall, the top predictors for having an ER visit differ from those that predict the number of ER visits. Of the top 10, there is only an overlap of two variables. Lower cost substance abuse and orthopedic trauma, fracture or dislocation, II are top predictors for both having an ER visit as well as the number of ER visits. Lower cost neurology and lower cost dermatology, are examples of strong predictors on having an ER visit but not for the number of ER visits.

Logistic Variable	Negative Binomial Count Variable
Orthopedic trauma, fracture or dislocation, II	Orthopedic trauma, fracture or dislocation, II
Orthopedic trauma, fracture or dislocation, III	19 to 34 year old females
Anti-emetic agents, not elsewhere classified, II OR Anti-emetic agents with noninsulin diabetic agents, selected steroids OR Anti-emetic agents, with CAD	Lower cost substance abuse
Lower cost substance abuse	19 to 34 year old males
Lower cost cardiology, I	Pct white - census variable
Member age	Pct high school - census variable
Low cost dermatology, I	Lower cost dermatology, II
Orthopedic trauma, fracture or dislocation, I	Asthma, COPD, II
Lower cost gastroenterology, I	35 to 44 year old females
Lower cost neurology	Hypertension, w significant complication/comorbidity

Table 3. Comparison of Logistic and Count Model Top Predictors

STEP FOUR: COMBINE MODELS FOR PREDICTED ER VISIT

With both regression models complete, the next step is to combine the output from both models to compute the predicted ER visit count. The Logistic Model output is the probability of any ER visit, P1. The Negative Binomial Model output gives the predicted ER visit count, pred_ER_count; however, since the model was run on the modified ER Visit count, you must add 1 to pred_ER_count. The predicted ER visit count is calculated as the following:

Member Predicted ER Visit Count = P1 * (1 + pred_ER_count)

STEP FIVE: MEMBER ER RISK SCORE

Each member now has a predicted ER visit count for the 12 month timeframe. The next step is to compute the member level ER risk score. The member ER risk score represents the member risk relative to entire population used to develop ER risk model:

Member ER Risk Score = Member Predicted ER Visit Count/ Book ER Visit Count

The Book ER Visit count is the average predicted ER visit count for members in the model. A member with a score of 1.0 has the same risk as the population, or in this case the book of business. A score of 0.90 indicates a 10% lower ER risk than the overall population.

APPLICATIONS OF MEMBER ER RISK SCORE

The member ER risk score now allows for risk adjusted ER utilization metrics. Table 4 shows unadjusted and risk adjusted ER metrics for providers. The table has the average ER risk score as well as unadjusted and risk adjusted average ER visit count and ER visits per thousand members. Unadjusted comparisons would have indicated that Provider 5 had the highest average ER visits with Provider 2 having the lowest.

The first record in the table demonstrates how the risk adjusted and unadjusted values are the same since the book of business average risk score is 1.0. You will notice that the members attributed to Provider 1 have the highest average ER risk score. When comparing Provider 1 to Provider 2, the unadjusted rates show that Provider 1 has more ER visits per thousand members. However, once you take into account that Provider 1 has a higher ER risk score; the risk adjusted rates bring the rates per thousand to almost the same as Provider 2. This risk adjustment is necessary because the patients attributed to Provider 1 have a 16% higher than average ER risk score, whereas Provider 2 only has a 5% higher than average ER risk score. In looking at Provider 5, the average risk score is slightly lower than average, so the risk adjustment slightly increases the rate. The risk adjusted metrics show that Provider 1 has the lowest ER visit count when you adjust for the fact that this Provider had the highest member risk pool.

ACO	Average ER Risk Score	Unadjusted Avg Number ER Visits	Unadjusted ER Visits/1000 members	Risk Adjusted Avg Number ER Visits	Risk Adjusted ER Visits/1000 Members
Book of Business	1.0000	0.1517	151.68	0.1517	151.68
Provider 1	1.1640	0.1317	131.70	0.1131	113.15
Provider 2	1.0526	0.1198	119.80	0.1138	113.81
Provider 3	1.0501	0.1224	122.43	0.1166	116.60
Provider 4	1.0008	0.1318	131.79	0.1317	131.68
Provider 5	0.9967	0.1535	153.48	0.1540	153.98

Table 4. Provider ER Utilization Metrics

CONCLUSION

This paper outlines the methodology used to model members' risk of ER visits. A two part retrospective regression model is used to predict ER usage which is then used to compute a member level ER risk score. The member ER Risk score is used to risk adjust utilization metrics to fairly compare providers based on the underlying member health risk of their patients.

The methodology presented was developed because member level risk scores should not be used to risk-adjust utilization metrics as member level risk scores are built to predict cost and not utilization. Results demonstrate that ER risk scores are not well correlated (r=0.323) with scores built to predict costs. Further, the member health conditions that are strong predictors of overall cost are not the same as factors for specific utilizations. Table 5 has a comparison of the top 10 predictors for the retrospective risk score model alongside the top 10 predictors for the two part ER risk model.

ERG Retrospective Risk	Logistic (ER Visit)	Negative Binomial Count (Number of ER Visits)
Ischemic heart disease, heart failure, cardiomyopathy, IV	Orthopedic trauma, fracture or dislocation, II	Orthopedic trauma, fracture or dislocation, II
Chronic renal failure, III	Orthopedic trauma, fracture or dislocation, III	19 to 34 year old females
Joint degeneration & major joint inflammation, III	Anti-emetic agents, not elsewhere classified, II OR Anti-emetic agents with noninsulin diabetic agents, selected steroids OR Anti-emetic agents, with CAD	Lower cost substance abuse
Malignant neoplasm, breast/female genital tract, with active mgmt, w/o significant complication/comorbidity, II	Lower cost substance abuse	19 to 34 year old males
Malignant neoplasm, gastroenterology, IV	Lower cost cardiology, I	Pct white - census variable
Joint degeneration & major joint inflammation, II	Member age	Pct high school - census variable
Malignant neoplasm, breast/female genital tract, with active mgmt, w/o significant complication/comorbidity, I	Low cost dermatology, I	Lower cost dermatology, II
Ischemic heart disease, heart failure, cardiomyopathy, VI	Orthopedic trauma, fracture or dislocation, I	Asthma, COPD, II
Normal pregnancy, delivery, I	Lower cost gastroenterology, I	35 to 44 year old females
Neoplastic blood diseases & leukemia, IV	Lower cost neurology	Hypertension, with significant complication/comorbidity

Table 5. Top 10 Predictors by Model

Application of the ER risk scores as adjustment factors for a practice or ACOs makes the performance metric more fair since it can deflate or inflate the metric to more appropriately capture the utilization that is driven by the patient's medical conditions that may predispose them to a need for emergent care in ER.

ACKNOWLEDGMENTS

Thank you to Nathan Hamilton, Laurel Trantham and Margarita Lankford for their collaboration and insight while building this two part regression model.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Tracy B. Lewis
Blue Cross and Blue Shield of North Carolina (BCBSNC)
919.765-3831
Tracy.Lewis@bcbsnc.com