

# How will EDGE affect your 2019 ACA Risk Adjustment Transfer?

Zachary M. Davis, FSA, MAAA

Phil Ellenberg, MS

Brian A. Sweatman, FSA, MAAA



For the first time beginning in 2019, the Centers for Medicare and Medicaid Services (CMS) will be calibrating the HHS-HCC commercial risk adjustment model—at least in part—using actual Patient Protection and Affordable Care Act (ACA) experience from the 2016 EDGE server data submissions. Up until the 2019 benefit year, CMS has based the model solely on non-ACA data.

This paper and the accompanying interactive exhibits allow the reader to review the coefficients from the 2019 model and compare how the EDGE data incorporated into the 2019 model will affect risk scores (which have a direct impact on an issuer's risk adjustment transfer). While future (2020 and later) risk adjustment models are unknown, it would not be unreasonable to assume that the weight assigned to EDGE data in creating the coefficients will increase; therefore, it would be prudent for ACA issuers to begin investigating how these model changes may influence their overall financial performance.

To demonstrate the potential impacts from the new 2019 coefficients, ACA issuers would need demographic and condition prevalence data for their population(s). This would allow for an estimation of how their plan liability risk score (PLRS) might change from one period to the next. However, in order to understand the resultant impact to their estimated risk adjustment transfer(s), issuers would also need prevalence data for the total market(s) in which they operate. For the purpose of providing a hypothetical PLRS impact, we created a sample population using over 1.9 million individual ACA members from Milliman's 2016 Consolidated Health Cost Guidelines Sources Database (CHSD). The development of this population is described in more detail in the methodology section below.

Before diving into the EDGE data coefficients, it's important to provide some background on the development of the model coefficients. The original 2014 risk adjustment model was calibrated using the 2010 Thomas Reuters (now IBM Watson) MarketScan (MarketScan) data set. This data set contained over 45 million members in all 50 states and the District of Columbia.<sup>1</sup> At the time the first risk adjustment model was being developed, CMS believed this was the best available data set. The advantage

of MarketScan was its large volume and broad cross section of national data, while the main disadvantage was that its underlying population, mostly employer groups, did not represent the characteristics of the ACA small group or individual populations. In order to calculate the risk adjustment transfer each year, CMS collects data through EDGE servers established by each ACA health insurance issuer. CMS has been collecting this data since the ACA market reforms were implemented starting in January 2014. One response to comments in the 2017 Notice of Benefit and Payment Parameters (NBPP) indicated that CMS did not collect enrollee-level data in the EDGE data collection, but intended to discuss incorporating this data in the future. The following year, CMS finalized its decision to collect de-identified enrollee-level data in the 2018 NBPP specifically for the purposes of recalibrating the risk adjustment model, informing development of the Actuarial Value (AV) Calculator and methodology, and calibrating other HHS programs in the individual and small group markets.<sup>2</sup> The decision to collect this enrollee-level data has paved the way for the incorporation of the 2016 EDGE data into the 2019 risk adjustment model.

Since the 2016 model year, CMS has used a methodology of calibrating three separate risk adjustment models using three separate years of MarketScan data and then applying an equal blend of each model's coefficients to create the final model for a particular year. For 2019, it is continuing this methodology, but using two years of MarketScan data (2014 and 2015) along with the 2016 EDGE data. In the past, the final blended coefficients have been published in the NBPP with no way to distinguish the coefficients from each separately calibrated model. Since the 2016 EDGE data was not ready when the draft 2019 NBPP was released, the draft notice only included the coefficients for the blended 2014 and 2015 models. When the final 2019 coefficients were published, we were able to reverse-engineer the EDGE components of the coefficients. The analytical development process CMS used to process the 2016 EDGE data is described in more detail in the methodology section.

1 "Risk Adjustment Methodology Overview." Department of Health and Human Services. Centers for Medicare & Medicaid Services. Center for Consumer Information and Insurance Oversight. May 21-23, 2012. Retrieved on August 2, 2018. <https://www.cms.gov/CCIIO/Resources/Presentations/Downloads/hie-risk-adjustment-methodology.pdf>.

2 "Enrollee-level EDGE Dataset for Research Requests." Department of Health and Human Services. Centers for Medicare & Medicaid Services. Center for Consumer Information and Insurance Oversight. May 18, 2018. Retrieved on August 2, 2018. <https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Enrollee-level-EDGE-Dataset-for-Research-Requests-05-18-18.pdf>.

The rest of this paper contains four interactive exhibits for the reader to compare the 2019 draft, final, and EDGE coefficients. To clarify, the draft factors represent the equal blending of the 2014 and 2015 calibrated MarketScan models, the EDGE factors represent the model calibrated to the 2016 EDGE data, and the final factors are the blend of the coefficients giving equal weight to the 2014 MarketScan, 2015 MarketScan, and 2016 EDGE data.

**Note that all data in the interactive exhibits have been updated to reflect the changes to the EDGE coefficients announced by CMS on July 27, 2018.**

Interactive 1 (interactives are best viewed in the full screen mode by clicking the arrow in the bottom right) is a table with each of the three sets of coefficients for all combinations of models (adult, child, and infant) and plan metal level. The user can toggle between metal and model and compare the differences between coefficients. There are five metal levels (catastrophic, bronze, silver, gold, and platinum) and three models (adult, child, and infant). The adult, child, and infant models are all structurally different. The risk score for an adult member is based on the sum of demographic, enrollment duration, diagnosis, prescription drug, and interaction factors. The risk score for the child model only includes demographic and diagnosis factors. Finally, the infant model includes a demographic factor along with 25 combinations of maturity and severity. For all models, the factors are additive. Note that no additional adjustments were made to these coefficients for the cost-sharing reduction (CSR) eligible members.

**FIGURE 1: HHS-HCC RISK ADJUSTMENT MODEL – DRAFT, FINAL, AND EDGE COEFFICIENTS**

Factor Description	Draft	Final	Edge
HHS_HCC066 (Hemophilia)	43.412	52.205	69.791
RXC_02 (Anti-Hepatitis C (HCV) Agents)	41.357	38.868	33.890
RXC_10_x_HCC159_158 (Additional effect for enrollees with Rx 10 (Cystic Fibrosis Agents) and (HCC 159 (Cystic Fibrosis) or 158 (Lung Transplant Status/Complications)))	29.949	29.406	28.320
HHS_HCC184 (End Stage Renal Disease)	31.086	29.491	26.301
G14 (Heart Assistive Device, Artificial Heart and Heart Transplant)	27.268	28.087	29.725
HHS_HCC041 (Intestine Transplant Status/Complications)	29.062	28.120	26.236
HHS_HCC125 (Respirator Dependence/Tracheostomy Status)	27.016	26.538	25.582
HHS_HCC158 (Lung Transplant Status/Complications)	25.270	25.358	25.534
HHS_HCC251 (Stem Cell, Including Bone Marrow, Transplant Status/Complications)	23.520	24.303	25.869
HHS_HCC008 (Metastatic Cancer)	20.694	21.018	21.666
RXC_08 (Multiple Sclerosis Agents)	19.754	19.734	19.694
HHS_HCC159 (Cystic Fibrosis)	10.767	13.669	19.473
RXC_10 (Cystic Fibrosis Agents)	10.149	12.461	17.085
G06 (Disorders of Bone Marrow)	11.123	12.232	14.450
RXC_09 (Immune Suppressants and Immunomodulators)	11.895	12.224	12.882
HHS_HCC009 (Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia)	11.689	11.829	12.109
HHS_HCC023 (Protein-Calorie Malnutrition)	11.365	11.404	11.482
HHS_HCC034 (Liver Transplant Status/Complications)	9.866	10.353	11.327
G10 (Quadriplegia and Traumatic Complete Lesion Cervical Spinal Cord)	8.834	9.744	11.564
HHS_HCC042 (Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis)	9.302	9.308	9.320
HHS_HCC153 (Atherosclerosis of the Extremities with Ulceration or Gangrene)	9.004	8.906	8.710
HHS_HCC226 (Hip Fractures and Pathological Vertebral or Humerus Fractures)	8.291	8.429	8.705
HHS_HCC253 (Artificial Openings for Feeding or Elimination)	8.005	8.129	8.377
HHS_HCC122 (Non-Traumatic Coma, and Brain Compression/Anoxic Damage)	7.866	8.062	8.454
HHS_HCC118 (Multiple Sclerosis)	7.959	7.932	8.606
SEVERE_V3_x_G06 (Severe Illness x HCC group G06 (G06 is HCC Group 6 which includes the following HCCs in the blood disease category: 67, 68))	8.532	8.038	7.050
SEVERE_V3_x_G08 (Severe Illness x HCC group G08 (G08 is HCC Group 8 which includes the following HCCs in the blood disease category: 73, 74))	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC006 (Severe Illness x Opportunistic Infections)	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC008 (Severe Illness x Metastatic Cancer)	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC009 (Severe Illness x Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia)	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC010 (Severe Illness x Non-Hodgkin's Lymphomas and Other Cancers and Tumors)	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC115 (Severe Illness x Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy)	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC135 (Severe Illness x Heart Infection/Inflammation, Except Rheumatic)	8.532	8.038	7.050
SEVERE_V3_x_HHS_HCC145 (Severe Illness x Intracranial Hemorrhage)	8.532	8.038	7.050
G07 (Diseases of the Blood)	7.217	7.622	8.432

To experience the full interactive nature of this exhibit, please visit: [tinyurl.com/EDGE-ACA-2019](https://tinyurl.com/EDGE-ACA-2019)

Interactive 2 is a summary of the model performance calculated and published by CMS from 2016 to 2019 as measured by each model's coefficient of determination or R<sup>2</sup> (a model's R<sup>2</sup> statistic represents the proportion of variance within a response variable explained by the model's explanatory variables). Each year starting with the 2016 model year, CMS published the R<sup>2</sup> statistic for each calibrated model. For the 2018 model year, CMS only published the R<sup>2</sup> statistic for the 2013 and 2014 data years. The user can toggle between metal and model to compare the historical results. The 2019 model year is the best performing model using the R<sup>2</sup> metric with the adult models performing the best followed by the infant and then the child model. One thing CMS did not clarify in the NBPP was the mixture of small group and individual EDGE data used in the calibration. Splitting out the individual and small group EDGE data and calibrating two separate models could be a consideration in the future to improve the model's ability to predict the liability of the underlying populations better than a blended model, although that would entail significant additional operational complexity for both CMS and issuers.

**FIGURE 2: HHS-HCC RISK ADJUSTMENT MODEL – HISTORICAL R<sup>2</sup> VALUES**



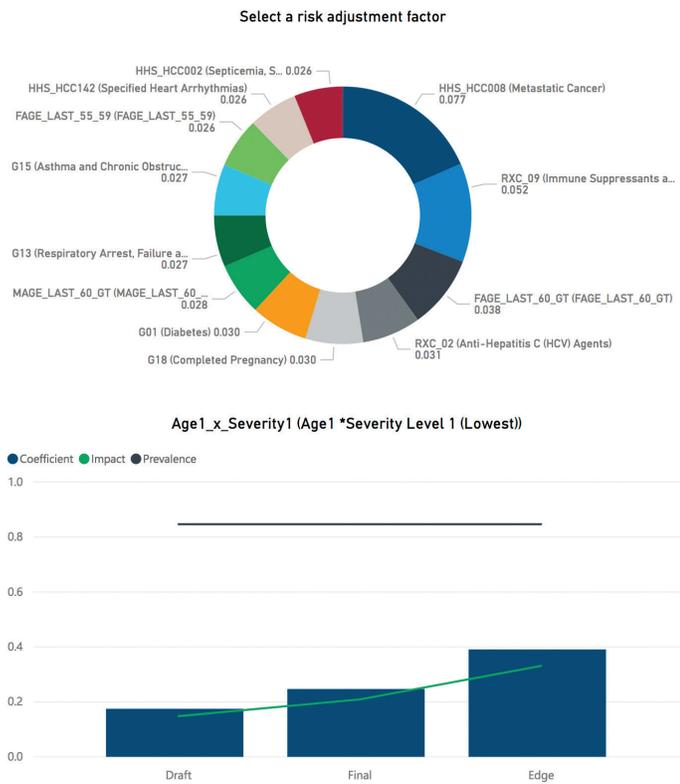
To experience the full interactive nature of this exhibit, please visit: [tinyurl.com/EDGE-ACA-2019](https://tinyurl.com/EDGE-ACA-2019)

In 2016, the Society of Actuaries (SOA) updated its study estimating the accuracy of over 40 different risk score models.<sup>3</sup> Overall, the HHS-HCC risk adjustment model using diagnoses only (the study was developed before HHS added prescription drugs for the 2018 model year) performed similarly to other models, but was noticeably below the best-in-class models. It is interesting to note that the incorporation of the EDGE data seems to have resulted in only modest improvements over the 2018 model year. Another interpretation is that the EDGE data performed well given the challenges issuers faced in the early years of EDGE data submission. Regardless of the interpretation, CMS could publish other useful statistics (such as the mean absolute percentage error, 'MAPE') that aid relative model comparisons.

3 "Accuracy of Claims-Based Risk Scoring Models." Society of Actuaries. October 2016. Retrieved on August 2, 2018. <https://www.soa.org/Files/Research/research-2016-accuracy-claims-based-risk-scoring-models.pdf>.

Interactive 3 allows the user to compare and contrast the 2019 final, 2019 draft, and 2019 EDGE coefficients for each indicator variable that was included in each model (adult, child, and infant). There are some large changes to the coefficients when comparing these three models; and in order to understand the aggregate magnitude of these changes prevalence rates were developed. As described in the methodology section below, we created a sample ACA population with over 1.9 million members from the Milliman CHSD to develop prevalence rates for each variable in each model. Using the prevalence multiplied by the 2019 coefficients, we are able to estimate an impact attributable to each component. Having this impact helps compare and contrast the three sets of coefficients. HCCs that have very high coefficients and low prevalence along with HCCs with low coefficients and high prevalence will have the largest impact on the final risk scores. Note that there are no HCCs that have both high prevalence and high coefficients. (Interactive 3 allows the user to right-click on a component and drill through for more detail on that component’s impact for each model.)

**FIGURE 3: HHS-HCC RISK ADJUSTMENT MODEL – DISTRIBUTION OF RISK ADJUSTMENT IMPACT**



To experience the full interactive nature of this exhibit, please visit: [tinyurl.com/EDGE-ACA-2019](http://tinyurl.com/EDGE-ACA-2019)

When the user drills through to a specific component, they can hover over the bar graph to compare the “impact” between the models. For example, when focusing on adult silver, G01 (diabetes)

shows an impact of 0.027 for the EDGE model and 0.031 for the final model. Taking this difference represents the potential change to an issuer’s risk score from this component. It is interesting to note that the demographic factors have the highest differences between the final and EDGE with the EDGE demographic factors showing much lower impacts compared to the final.

Looking at the infant silver model, the EDGE components have a positive impact on 19 of the 27 components (2 age/gender components and the 25 combinations of maturity and severity) and these will have a significant increase on the overall infant risk score. Any carriers with significantly larger or smaller infant populations could see large changes to their risk scores. The user can also see this in Interactive 4.

Interactive 4 allows the user to compare the components of the PLRS for each combination of model and metal.

**FIGURE 4: HHS-HCC RISK ADJUSTMENT – TOTAL SIMULATED RISK SCORES**

Metal	Draft	Final	Edge
Catastrophic	1.186	1.142	1.053
Bronze	1.192	1.152	1.072
Silver	1.322	1.276	1.183
Gold	1.466	1.413	1.306
Platinum	1.607	1.550	1.436

To experience the full interactive nature of this exhibit, please visit: [tinyurl.com/EDGE-ACA-2019](http://tinyurl.com/EDGE-ACA-2019)

In the adult model, the total HCCs and RXCs are relatively consistent for the sample population with the HCCs component being higher for the EDGE and RXC being higher for the draft and final versions. The demographic components of the EDGE models are significantly lower. This could negatively affect carriers that have healthier than average membership because a majority of members will be receiving no HCC or RXC—only the duration and demographic components. Conversely, it could have a positive impact on carriers that have riskier members as more weight will be put on the HCC and RXC (diagnosis and drug utilization) components of the risk scores. In the child model, the EDGE risk scores were slightly higher for the demographic and HCC components. In the infant model, the EDGE risk scores are significantly higher (mainly driven by the change to the Age1 \* Severity Level 1 component).

## Methodology

To simulate the prevalence rates of the various risk adjustment factors among ACA members, we utilized Milliman’s 2016 CHSD. We first identified all members in the CHSD who have individual ACA coverage. This totaled to over 1.9 million members nationwide. We then compiled demographics, enrollment information, medical claims, and prescription claims for these members. This information was subsequently processed through the 2018 HHS-Developed Risk Adjustment

Model Algorithm “Do It yourself (DIY)” software published by CMS to develop prevalence rates for each component of the model. The CMS model assigned HCC, RXC, and other risk adjustment factors to each member according to CMS’s algorithms. We used the output from the CMS model to determine the per capita prevalence rate of each risk adjustment factor within the CHSD population. Because the 2019 model has not been published by CMS, we used the 2018 model for development of the prevalence rates and applied those to the 2019 coefficients to calculate the final impacts.

The 2019 EDGE model coefficients were developed by taking three times the final 2019 coefficients published in the final 2019 NBPP minus two times the coefficients published in the draft 2019 NBPP. After applying this methodology, a very small number of coefficients are slightly negative. In order to keep all the data in a replicable format, no adjustments were made to these coefficients. We believe these negative coefficients are most likely due to rounding. One key assumption in our analysis is that the draft coefficients based on the 2014 and 2015 MarketScan data were used unchanged in the final coefficient blending with the EDGE data. If those coefficients changed materially in the final model, it would affect these results.

The analytical development process used on the 2016 EDGE data was described in the Final 2019 NBPP: *“We arrived at the 2016 enrollee-level EDGE analytical dataset using several criteria. We limited the sample to ages 0–64 to maintain the same age categories as those HHS has used in the MarketScan data, with which the EDGE coefficients are blended. Currently, we use the age 60–64 factors for those over 65 years of age enrolled in individual and small group market coverage, and will continue to do so for the 2019 benefit year. We will consider whether to propose expanding the age and sex factors to include age groups and associated costs for enrollees ages 65 and above in future model recalibrations. We also excluded derived claims, any newborn diagnoses for infants older than one year of age, anomalous claims (for example, pregnancy diagnoses if sex is male) and those with sex unknown. There were approximately 47 million, 28 million and 31 million total unique enrollees in the 2014 MarketScan, 2015 MarketScan, and 2016 enrollee-level EDGE data, respectively. Relative risks were similar in the 2016 enrollee-level EDGE data for most categories in all three adult, infant and child samples. As mentioned above, enrollee-level EDGE data reflected lower spending and relative risk patterns for shorter enrollment duration enrollees compared to MarketScan data.”*<sup>4</sup>

4 “Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2019.” Federal Register, Vol. 83, No. 74, page 16941. April 17, 2018. Retrieved on August 2, 2018. <https://www.gpo.gov/fdsys/pkg/FR-2018-04-17/pdf/2018-07355.pdf> (page 16941).

## Caveats and Limitations

Brian Sweatman and Zach Davis are Members of the American Academy of Actuaries and Fellows of the Society of Actuaries and meet the Qualification Standards of the American Academy of Actuaries.

The values shown here are based on the average of the underlying CHSD data set. Results for any particular stakeholder may vary from those presented here due to but not limited to different underlying populations and future changes to laws and regulations.

Note that the CHSD is 2016 data, whereas the ‘DIY’ model provided by CMS is a 2018 model. Therefore, any major differences in ICD-10 mapping between these two time periods may impact diagnoses with significant coding changes.

### FOR MORE ON MILLIMAN’S PERSPECTIVE ON HEALTHCARE:

Visit [milliman.com/insight/healthcare](http://milliman.com/insight/healthcare)

Visit our blog at [healthcaretownhall.com](http://healthcaretownhall.com)

Follow us at [twitter.com/millimanhealth](https://twitter.com/millimanhealth)

---

### CONTACT

Zachary M. Davis  
[zach.davis@milliman.com](mailto:zach.davis@milliman.com)

Phil Ellenberg  
[phil.ellenberg@milliman.com](mailto:phil.ellenberg@milliman.com)

Brian A. Sweatman  
[brian.sweatman@milliman.com](mailto:brian.sweatman@milliman.com)