

Estimating ACA risk transfers: It's a tough job, but someone has to do it

Peter Fielek, FSA, MAAA

Erik Huth, FSA, MAAA

Jason Petroske, FSA, MAAA

If you're responsible for estimating Patient Protection and Affordable Care Act (ACA) risk adjustment transfers, you're in the right place.

You probably understand how material risk transfers can be to the bottom line¹ and are familiar with how extensively they are integrated into key processes, including year-end financial statements, experience analyses, forecasting, and annual pricing exercises. The problem? Making a reliable estimate is challenging—the calculation can be confusing, its pieces are full of uncertainty, and the estimate must occur well before the Centers for Medicare and Medicaid Services (CMS) releases results the following summer. Given all this, it wouldn't be surprising if you've had to explain why the prior year's results missed forecast and why this year's projections will be more accurate. This paper may help address both questions.

The transfer calculation contains a number of components, both for the individual carrier and for the entire state. The following represent the major items you'll need to consider:²

Component	Description	Applicability	
		Carrier	State
Plan liability risk score (risk score)	Measurement of the population morbidity level	X	X
Allowable rating factor (ARF)	Representation of the population average age	X	X
Actuarial value (AV)	Measurement of plan benefit richness	X	X
Induced demand factor (IDF)	Measurement of plan benefit richness	X	X
Geographic cost factor (GCF)³	Reflection of the average cost of healthcare in a given region	X	
State Average Premium	Reflection of the average cost of healthcare across the state		X
Billable Member Months (BMMs)	Representation of a carrier's market exposure	X	

With this long list of items, the lack of readily available competitor data, and, sometimes, several additional and impactful “unknowns” (such as major risk adjustment model changes or unexpected healthcare utilization patterns caused by a worldwide pandemic), the barriers surrounding a reasonable transfer estimate can seem insurmountable.

Thankfully, all is not lost. There are helpful data sources and methods to lessen some of the uncertainty. By the end of this paper, we aim to provide insights and options for estimating key risk adjustment components and increasing the level of confidence in your projections.

Start with what's handy

All items in the prior table *must be* addressed, and a practical strategy would be to start where the most data is available—the *carrier's own* experience. For instance:

- The ARF, AV, and IDF can be calculated from calendar year plan and enrollment data and the codified federal and state values.⁴ Even with a few months remaining, these factors are usually very stable, changing little by year-end.
- The risk score should reflect the latest claim and enrollment data—either from an enterprise data warehouse or an External Data Gathering Environment (EDGE) server. For incomplete data, a factor will be needed to bring the risk score up to a basis consistent with EDGE data at the submission deadline (incurred through December and paid through the following April). This factor might be based on either prior experience or projections of future performance and could account for additional claim incurrals, possible EDGE risk score logic changes, and future supplemental diagnoses submissions.

When using prior year completion to inform future trends, it might be helpful to calculate the completion level at several points in time to get a better “feel” for the patterns. For carriers with smaller blocks of business, the relative completion variability will tend to be higher, so additional analysis may be necessary to develop the proper factor.

A carrier should first focus on establishing sound estimates for the values it knows and are within its control. Of course, any transfer projection will contain some level of uncertainty, but shoring up one side of the “equation” goes a long way in reducing the variability over time.

Moving into more unknown territory

The logjams begin when projecting the *statewide* factors for a given risk pool (i.e., individual, catastrophic, or small group). Some are more straightforward, often due to their relative stability year-over-year or the availability of useful published data. The next set of factors can be reasonably estimated with the least effort.

- Changes in the individual market ARF,⁵ AV, and IDF can be projected through plan selection and age trends over the prior year in open enrollment activity, as reported in the CMS public use files (PUFs). Because open enrollment data doesn't exist in the small group market and some state-based exchanges in the individual market, these factors may come, instead, from prior year trends implied from the annual risk adjustment reports.⁶
- Statewide average premium changes typically follow closely from approved ACA rate changes,⁷ which CMS publishes in the Uniform Rate Review Template (URRT) PUFs. Any calculated rate change from the PUF may need to be adjusted for benefit buy-downs, market movements, and differences between historical premium changes versus carrier-estimated rate changes.⁸ Further, carriers need awareness of any extraordinary circumstances affecting the market that year, such as 2020 premium refunds precipitated from favorable claim experience in the wake of the COVID-19 pandemic.
- GCFs present greater difficulties because the values are dependent on cost changes within specific ACA regions—the lower relative membership and carrier-specific decisions create greater volatility than those factors dependent on the state in total. GCFs can mirror the prior year's values or can

be estimated in a more complex manner, such as by reviewing either published silver rate changes by rating area or implied area factors in URRT Worksheet III.

The above approaches are not perfect but are certainly useful when little market information exists.

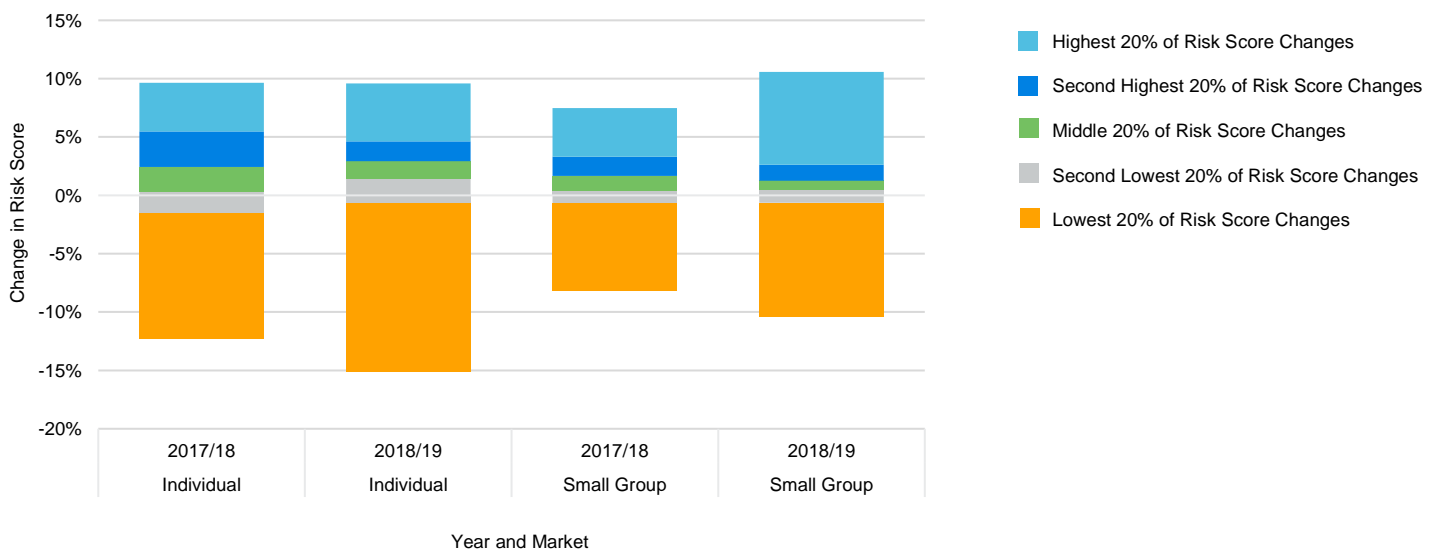
One more to go (but it's a big one)

Not surprisingly, the statewide risk score is, by far, the most difficult factor to project. Not only is the current year's data virtually nonexistent, but the carrier level risk scores themselves can change significantly each year. Figure 1 shows the variability in the statewide risk score change by market the past two years, separated by quintile.

For example, the change in the 2019 statewide individual market risk score varied by state, from a 15.2% decrease and a 9.6% increase (i.e., an absolute change of 24.8 percentage points). Further, the 10 states with the smallest 2019 risk score change (in orange) ranged between a 15.2% decrease and a 0.7% decrease alone.

A 1.0% change in the statewide risk score can be worth an average of \$4 to \$6 per member per month (PMPM) in transfers. Meaning, a carrier with 100,000 member months that underestimates the statewide risk score by 4.0% could experience a \$1.6 million and \$2.4 million unfavorable restatement once results are known.⁹ In other words, the financial consequences of a poor risk score projection can be incredibly impactful, and the degree of annual variability in Figure 1 likely suggests material transfer estimation errors are rather commonplace.

FIGURE 1: CHANGE IN STATEWIDE RISK SCORE BY QUINTILE¹⁰



As with other statewide factors, there are several ways to approach the statewide risk score. We describe the most common below:

PARTICIPATE IN STATEWIDE TRANSFER STUDIES

In these studies, carriers contribute data to a third party in return for aggregated results with, presumably, credible insights into the market as a whole. For many reasons, adjustments to participant risk scores may be appropriate if the study:

- Does not represent 100% of the market.
- Does not reflect data as of the EDGE submission deadline.
- Does not include all known or anticipated supplemental diagnosis submissions.¹¹
- Uses data with known issues that will be eventually corrected.

While the information can be useful, arranging these types of studies can be complex, given the unique incentives for each carrier in the market. Results are strongest when there is sufficient participation—a “critical mass,” if you will—which means a large portion of carriers will need to share data. However, dominant carriers have historically been more protective of information that has tended to benefit others in the state more than themselves. With the onset of COVID-19, though, even those with high market share may find their size is not enough to insulate them from significant inaccuracies.

LEVERAGE THE CMS INTERIM RISK ADJUSTMENT REPORT

CMS typically releases a report in March with key risk adjustment metrics on EDGE servers as of a specific point in time.¹² The report can be appropriate if prior interim results show a reasonable level of accuracy. The late release date, however, likely means the interim report cannot be used for year-end financial reporting, but it may be sufficient for ACA pricing purposes.

Carriers should use the interim report with caution, though, as CMS requirements do not guarantee complete data accuracy or even data consistency among carriers—either within or across years.¹³ In fact, there’s no clear incentive for carriers to provide more data than is necessary (although many likely do). Further, the numbers in the reports do not represent final results, and carriers commonly submit additional claim and supplemental diagnosis files until the final days before the EDGE deadline.

DEVELOP THE RISK SCORE ESTIMATE “FROM SCRATCH”

For transfers needed before the interim report’s release, carriers without access to a risk adjustment study can leverage the previous year’s published risk score. Naturally, this is a starting point only, and adjustments should be made for items affecting the current year. We explain some of these next, using the individual market as an example, given the larger pool of available and relevant data compared with the small group market.

Changes to CMS risk adjustment methodology

Annually, CMS makes various changes to the risk-scoring methodology underlying the Health and Human Services Hierarchical Condition Category (HHS-HCC) model, ranging from recalibrations for more recently available data to material structural changes. The items CMS tends to adjust fall into three main changes:

1. Updates to age/gender factors

A carrier can approximate this impact by running its benefit year enrollment through both the current and previous year’s CMS risk-scoring algorithm. The change can then be applied to the market, assuming a similar distribution of membership by age, gender, and plan.

2. Updates to HCC coefficients and logic

A carrier can estimate this in ways similar to age/gender factor changes. However, the distribution of HCCs among carriers is often much more volatile than age/gender, depending on carrier size. As such, larger HCC coefficient impacts could mask the true results if the carrier’s HCC prevalence is different from the market. Access to a large, more credible data set may help reduce the variability.

Further, if CMS adds or removes HCCs or significantly changes the structure of the model, as in the 2018 and 2021 benefit years, then carriers will need to account for these potential impacts as well (possibly without full guidance from CMS on the final model logic).

3. Updates to prescription drug condition (RXC) coefficients and changes to the National Drug Code (NDC) list

This risk score impact is the most difficult to estimate in advance. A carrier cannot necessarily apply one, consistent data set to two different HHS-HCC models, as this method invariably ignores several key prescription drug utilization trends (which are generally more fluid than medical utilization trends):

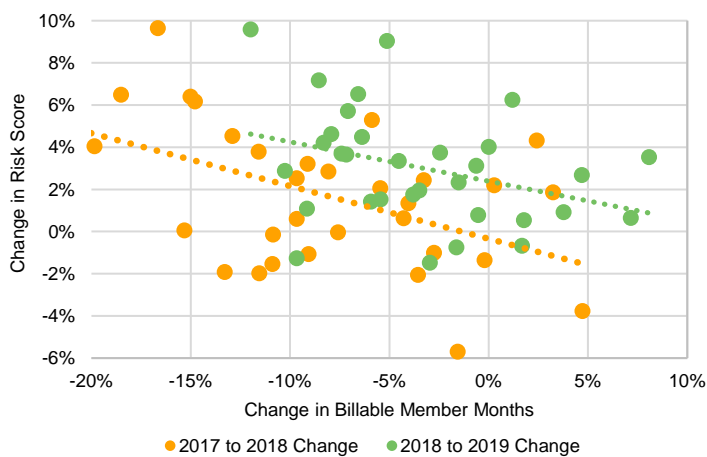
- a. Applying both the previous and current year methodology to the *previous year’s* experience ignores new prescription drug treatments now available, which will understate the current year’s risk score.
- b. Applying both the previous and current year methodology to the *current year’s* experience lowers the utilization of the previous year’s drugs that have been recently replaced, which will understate the previous year’s risk score.
- c. Applying the current year methodology to current experience and the previous year methodology to previous experience provides a better representation of the risk score but, unfortunately, will also include morbidity and enrollment changes that require *further* adjustments.

Aside from these considerations, CMS regularly updates the NDC list (and may update other logic) throughout the EDGE cycle. Some of these changes have been impactful, so the selected methodology or an explicit adjustment made at one point could be outdated by the final calculation at the end of April.

Changes in market enrollment

Enrollment increases tend to signal a decrease in the statewide risk score due to the addition of relatively healthier lives. Conversely, decreasing enrollment is typically an indicator of anti-selective lapsation, which leaves members with higher average risk scores in the market. Figure 2 highlights this relationship in the individual market for the prior two years.

FIGURE 2: CORRELATION BETWEEN THE CHANGE IN BILLABLE MEMBER MONTHS AND RISK SCORE IN THE INDIVIDUAL MARKET¹⁴



It's important to note, the correlation presented using actual CMS data is not perfect and can be obscured by other items influencing risk score movements, such as how the updated HHS-HCC model affects the state/markets' populations, medical coding trend, carrier/industry marketing efforts, and open enrollment timeframes, among others.

Another potential change in market enrollment relates to member shifts *among* issuers. When a large portion of enrollees shifts (such as when the carrier with the lowest premium rate suddenly changes), the new carriers lose known medical histories, which can depress their risk scores. Carriers could leverage open enrollment PUFs, which include the number of new consumers, total re-enrollees, active re-enrollees, and automatic re-enrollees. The file does not measure the number of re-enrollees switching health plans, but a carrier could estimate this by reviewing its own internal membership turnover or by reviewing historical changes in market share for ACA market participants.

Changes in the average ARF

An increase in ARF correlates to a higher risk score, both because the age/gender risk scores increase with age and because an older population will tend to have a higher prevalence of identified HCCs and RXCs. As stated earlier, carriers can reasonably approximate the change in statewide ARF but will have to estimate how much the statewide risk score rises for each percentage point increase in that ARF.

Changes in the average AV

Because risk scores reflect the carrier's liability, a decrease in AV (i.e., a shift to plans with leaner benefits) will generally lead to lower carrier claim costs and, therefore, is associated with a lower risk score in the risk adjustment model. Similarly to the ARF, carriers can make a decent estimate for the change in the statewide AV but will also have to translate that into a change in the statewide risk score.

Change in market claim utilization levels

Due to the COVID-19 pandemic, 2020 medical service utilization decreased significantly, generally between mid-March and the end of May, with total 2020 claim utilization decreasing overall (i.e., lower first-half claim utilization was not fully offset by higher second-half claim utilization).¹⁵ While substantial uncertainty remains, lower 2020 claim utilization will reduce 2020 risk scores, and carriers will have to account for impacts such as this to the statewide risk score.

Change in medical chart review efforts

Medical coding patterns differ not only by state but also by regions within a state. Coding accuracy and completeness start with the providers, but the metrics can be affected by carriers—especially those with robust chart review and outreach efforts. Even with all carriers striving to reach the same point, the level of coding will vary, as highlighted by the Risk Adjustment Data Validation (RADV) audit results. As carriers improve their coding over time, the statewide risk score will change to reflect any progress.¹⁶

Unfortunately, there is no available information to gauge coding trend in a given market. For instance, some states experienced large RADV adjustments in back-to-back years, while other adjustments dropped significantly or disappeared altogether. In fact, more than half of the non-catastrophic individual and small group markets (58 of 98) had some RADV impact affecting either the 2018 or the 2019 risk score.¹⁷

Going forward, it's reasonable to assume coding trend will begin to approach a more stable long-term level, but much of that is contingent on the current level of coding and the rate at which carriers target this area for improved risk adjustment performance.

Know what you know; know what you don't know

The sections above develop a foundation for estimating a change to all relevant components in the risk transfer equation,¹⁸ but it is not an exhaustive list. And, correctly projecting each item will not remove all the uncertainty—random claim and enrollment fluctuation will always be a wild card. Given this, it's prudent for carriers to hold some level of margin after developing a best transfer estimate, especially when incorporating the amounts into financial statements. There are many options for adding margin, and the following three methods may be appropriate depending on the situation:

- Add a fixed PMPM to a transfer payment (or subtract from a transfer receipt) and multiply the new amount by the BMMs.
- Increase the statewide risk score by some factor (say 1.05) and recalculate the transfer.
- Develop a range around *each assumption* in the equation and, either deterministically or through a simulation, recalculate transfer amounts.

Each of these methods will likely provide different results but can help inform an appropriate range of margin levels. It's important to note, it is *inappropriate* to simply apply a factor directly to the total transfer, as is typically done when setting a claim reserve. This approach will not set margin commensurate with the level of risk (i.e., a large and a small carrier may both add 10% margin to a \$1 million risk adjustment payable, but if each misses its estimate by a similar PMPM amount, the larger carrier will likely miss by much more than the \$100,000 margin held).

By utilizing the approaches outlined in this paper, it's possible to improve the accuracy, reliability, and usefulness of your projections or create further data points to compare with other methods. And, hopefully, the next time you're facing a barrage of questions about your ACA risk adjustment transfer estimates, you now have the tools and insights to present a strong case for why they can be so difficult to calculate and how you plan to develop a more robust process going forward.

Caveats, limitations, and qualifications

Peter Fielek, Erik Huth, and Jason Petroske are actuaries for Milliman and members of the American Academy of Actuaries. They meet the qualification standards of the Academy to render the actuarial opinion contained herein. To the best of their knowledge and belief, this paper is complete and accurate and has been prepared in accordance with generally recognized and accepted actuarial principles and practices.

The material in this paper represents the opinion of the authors and is not representative of the views of Milliman. As such, Milliman is not advocating for, or endorsing, any specific views contained in this paper related to the ACA risk adjustment program.

The information in this paper is designed to provide a framework for developing CMS risk adjustment transfer assumptions. This information may not be appropriate, and should not be used, for other purposes. The authors do not intend this information to benefit any third party that receives this work product. Any third-party recipient of this paper that desires professional guidance should not rely upon Milliman's work product but should engage qualified professionals for advice appropriate to its specific needs. Any releases of this paper to a third party should be in its entirety.

In preparing this analysis, the authors relied upon publicly available information from CMS, which they accepted without audit. However, they did review it for general reasonableness. If this information is inaccurate or incomplete, or at any time materially changes, the conclusions drawn from it may also change. Actual risk adjustment factor values will differ from those developed based on the methodology described in this paper, for the many reasons highlighted.

CONTACT

Peter Fielek
peter.fielek@milliman.com

Erik Huth
erik.huth@milliman.com

Jason Petroske
jason.petroske@milliman.com

ENDNOTES

- ¹ For benefit year 2019, CMS reported a variation in the individual (non-catastrophic) risk pool of between 13% of premium payment and 16% of premium receipt for carriers in the bottom and top quartile, respectively, of experienced claims. This range of variation still does not reflect the largest value any one carrier will have experienced in 2019. See <https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/RA-Report-BY2019.pdf>.
- ² The most current risk adjustment formula methodology is described in the 2021 Benefit and Payment Parameters (beginning with printed page 29192). See <https://www.federalregister.gov/documents/2020/05/14/2020-10045/patient-protection-and-affordable-care-act-hhs-notice-of-benefit-and-payment-parameters-for-2021>.
- ³ Although typically the state composite factor is not a 1.00 value, it is typical to assume 1.00 in the transfer calculation.
- ⁴ For instance, the AV and IDF for silver plans are 0.70 and 1.03, respectively, while the ARF values vary by state.
- ⁵ In 2018, CMS changed the federal age factors for members under age 20, which significantly changed the ARF in 2018 over 2017 in all markets.
- ⁶ This is possible because the small group results have remained reasonably stable year over year. Other than the change in the 2018 ARF, the majority of states and markets experience less than a 0.005 annual change in the AV and IDF and less than a 0.010 annual change in the ARF.
- ⁷ Approved rate filings can be found in www.ratereview.healthcare.gov.
- ⁸ For example, the composite rate increase calculation for a state and market needs to include a weight, which can be based on carrier-reported historical membership counts or carrier-projected membership counts—each of which will very likely deviate from the true enrollment in the benefit year being calculated. In the small group market, quarterly rate filings need consideration as well.
- ⁹ Assuming all other risk adjustment factor estimates are correct.
- ¹⁰ By definition, each color band represents one-fifth of the states (10 total).
- ¹¹ Typically from medical chart reviews but they could also correct diagnosis information from a source system in lieu of claim reprocessing.
- ¹² Historically at some point in January following the benefit year.
- ¹³ CMS does require certain data completeness thresholds, but, because the standard is not 100% accuracy, additional variability is introduced into the results.
- ¹⁴ We highlight the individual market only and exclude states with fewer than 1 million member months in 2019 to decrease the volatility in the presented results. We further remove Virginia due to large enrollment decreases (over 20%) in 2019. The patterns are similar for the individual market in total as well as the small group market.
- ¹⁵ Milliman MedInsight. Webinar: COVID-19 Impact on Healthcare Utilization, Enrollment, and Cost: Insights From Emerging Experience and Modeling. Retrieved January 26, 2021, from <https://www.medinsight.milliman.com/en/resources/covid19-impact-on-healthcare-utilization-enrollment-and-cost>.
- ¹⁶ Whittal, K. & Norris, D.T. (January 2016). Value of ACA Coding Improvement: Market Share and Market Effects. Milliman Healthcare Reform Briefing Report. Retrieved January 26, 2021, from https://us.milliman.com/-/media/milliman/importedfiles/uploadedfiles/insight/2016/2157hdp_20160127_value-of-aca-coding.ashx.
- ¹⁷ Rounding the RADV impact to the nearest 0.1%, 24 of 49 individual markets had no RADV adjustment in both 2018 and 2019, and 16 of 49 small group markets had no RADV adjustment in both 2018 and 2019. Review this article for additional information on RADV metrics: <https://us.milliman.com/-/media/milliman/pdfs/articles/a-look-behind-the-curtains.ashx>.
- ¹⁸ This does not directly account for RADV risk score adjustments or payments or recoveries from the high-cost risk pools. Ultimately, these items can have an equally impactful effect as any estimation errors in the base transfer calculation.