Provider coding accuracy in commercial value-based contracts

Jared Hirsch
Howard Kahn, FSA, MAAA
Rong Yi, PhD

Milliman

> As the prevalence of value-based contracts grows, so too does the variability of a provider's overall compensation.

In value-based contracting, providers are "measured" on certain performance metrics that will dictate how much they are ultimately paid. For example, it is common for a value-based contract to retrospectively compare the total actual annual spend of patients attributed to a provider entity to a predetermined spend target. This spend target is often risk-adjusted to account for any changes in the underlying attributed populations from year to year. If the actual annual spend is below the target, the provider entity shares in the "surplus"; otherwise, the provider entity is liable for a share of the "deficit." Therefore, risk adjustment plays an integral role in determining the provider's "measurement" of financial performance.

For the value-based contract to be equitable to all parties, risk adjustment should endeavor to accurately capture the change in the underlying population's morbidity to effectively measure the provider's true cost impact on its attributed population. *As we will illustrate, accurate data is critical for accurate risk adjustment, and improving risk adjustment accuracy can significantly change a provider's calculated measurement on performance metrics.*

In this paper, we explore the following aspects of risk adjustment as it pertains to value-based contracting:

- Identifying common claim data issues
- Quantifying the potential impact of rectifying these data issues
- Discussing how provider organizations can improve data accuracy

Risk adjustment has been fairly well established in Medicare Advantage and provides a basis for value-based contracting with Medicare payers. Because it is becoming increasingly more common for provider organizations to also enter into value-based contracts with commercial payers, our paper will focus on value-based contracting for commercial lines of business; however, we believe that our findings and conclusions are valuable to other lines of business and programs in which risk adjustment also plays a role.

Background on diagnosis-based risk adjustment in value-based contracts

Value-based contracts typically use diagnosis-based risk adjustment to measure the relative health status of a provider's attributed patient population. Diagnosis-based risk adjustment methodologies use standard administrative healthcare claims data (typically only medical claims) and membership data to assign relative risk scores at the patient level. A common risk adjustment methodology assigns each International Classification of Diseases (ICD) diagnosis code to a condition category (or identifies the code as not contributing to the risk score). Condition categories are intended to be clinically homogenous and statistically robust. Each condition category is assigned a relative risk weight to represent that category's relative contribution to overall cost. A patient's final risk score is calculated by summing all the relative risk weights associated with that individual's conditions. An individual's demographics (i.e., age and gender) may also have their own separate risk weights.

The risk score can be converted to a prediction of healthcare spending. For example, in a population where the average per member per month (PMPM) spending is $400, an individual with a 1.0 risk score is expected to cost $400 PMPM, and an individual with a risk score of 0.5 is expected to cost $200 PMPM. Taking an average of the risk scores from all individuals in a group, the group's expected healthcare spending can be compared to that of another group, making it feasible for direct cost comparisons among different patients, even with different health status and demographics. In value-based contracting, a provider's spend target is typically adjusted by the change in annual risk score to more accurately reflect the anticipated cost of the provider's current patients.

Common data issues that impact risk scores

An accurate risk adjustment methodology would result in a sicker (i.e., generally more costly) patient always having a higher risk score than a healthier (i.e., generally less costly) patient. In practice, the accuracy of risk adjustment is dependent on a provider's capability in capturing the most accurate diagnosis information on patients. Common issues relating to diagnosis coding and data submissions include:

1. **Not including all medical diagnoses on the claim.** The cause for missing medical diagnoses might stem from poor provider coding practices (such that the diagnosis is never captured) or a case where the diagnosis is captured but does not make its way to the claims feed that is used in the value-based contracting measurement calculation (i.e., technical issues with the data repository).

   Generally speaking, the capture of more diagnosis codes may result in higher risk scores. However, there are some important exceptions—not all types of claims count toward the risk scores. Typically, diagnosis codes on lab, radiology, durable medical equipment, and transportation claims are not used in risk scores. In addition, only unique diagnoses count toward the risk score. Multiple visits of the same medical diagnoses do not result in higher risk scores so as not to create incentives for overutilization of healthcare services.

2. **Not coding to the most appropriate level of precision.** The ICD-10 diagnosis codes can be three to seven characters long, in which the first one to three characters indicate the category of the diagnosis; the fourth to sixth characters indicate etiology, anatomic site, severity, or other clinical detail; and the seventh character indicates extension (such as initial or subsequent counter, etc.). More specific coding, i.e., using more characters to describe the illness, could help with identifying higher severity of the same diagnosis type, which in turn might lead to a higher, more accurate risk score.

3. **Not coding chronic conditions persistently year over year.** In our experience, chronic conditions that are highly prevalent in a commercial population, such as asthma and diabetes, may be coded one year and yet inadvertently not show up again the following year. A chronic condition typically does not resolve itself within a year, leading to the conclusion that it is very likely that care was given to the patient but the provider did not accurately capture it in the claims submission, leading to an eventual underestimation of that patient's risk score.

4. **Missing medical diagnoses suggested by the presence of prescription drug data.** Not all drugs are direct indications of conditions, and some drugs may be used off-label or used to treat conditions other than the primary indicator. However, there are some drug classes that are exclusively used for specific medical conditions.

Generally, improvements in the completeness and accuracy of coding may result in an increase or decrease in risk score. In the section below, we will explicitly quantify the potential change in risk score resulting from rectifying common issues #3 and #4 above.

Quantifying risk score improvement opportunities

Using 2014 and 2015 Massachusetts[1] claims data from the Truven Health Analytics MarketScan[2] (MarketScan) Commercial Database and Milliman's Contributor Health Source Database[3] (CHSD), we examined the extent to which the lack of coding persistency in chronic conditions (common issue #3) and the lack of medical diagnoses corresponding to prescription drugs (common issue #4) may impact risk scores.

For this analysis, we applied the 2015 U.S. Department of Health and Human Services Hierarchical Condition Category (HHS-HCC) risk adjustment model,[4] which is used to transfer funds between payers in the individual and small group markets. The 2015 HHS-HCC risk adjustment model classifies medical diagnosis codes, both ICD-9 and ICD-10, into more than 120 hierarchical condition categories (HCCs). We chose the HHS-HCC model because it is open source and has design features that are similar to most risk adjustment methodologies used in provider alternative payment contracts. The actual results of this study may vary if a different risk adjuster is chosen.

---

1   For illustrative purposes, we chose one state to run our analysis on. Commercial value-based contracts are more prevalent in Massachusetts than most other states, which would suggest that our results may be magnified in other states.

2   The MarketScan database contains all paid claims generated by approximately 35 million commercially insured lives. The MarketScan database represents the inpatient and outpatient healthcare service use of individuals nationwide who are covered by the benefit plans of large employers, health plans, government, and public organizations. The MarketScan database links paid claims and encounter data to detailed patient information across sites and types of providers, and over time. The annual medical database includes private sector health data from approximately 100 payers.

3   The Milliman CHSD provides enrollment and claims data contributed by health plans. The contributors grant Milliman permission to use the data research. The data are for commercial, individual, Medicare Advantage, Medicare Supplement, and Medicaid members.

4   HHS. HHS-Developed Risk Adjustment Model Algorithm "Do It Yourself (DIY)" Software Instructions. Retrieved September 6, 2017, from https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/DIY-instructions-10-16-15.pdf.

Chronic conditions: Persistency opportunity

We selected nine common chronic HCCs in a commercial population, examined their coding persistency rates over a two-year period in the study data set (2014 and 2015), and estimated the impact on risk score if coding persistency improves. The nine HCCs were selected because they have meaningful prevalence rates in the data, significant risk weights in the HHS-HCC model, and are at the highest hierarchy in the HHS-HCC clinical classification, meaning that patients are not expected to migrate to different HCCs within the same clinical hierarchy over time, which makes the persistency calculations more straightforward.

To estimate the risk score impact resulting from improved coding to fill in the gaps of chronic condition persistency year over year for these nine HCCS, we used the following steps:

1. Limited the study population to members with continuous enrollment in both Year 1 (2014) and Year 2 (2015).

2. Ran the 2015 HHS-HCC risk adjustment model on both years.

3. Identified potential gaps in chronic condition persistency by flagging patients with one of the nine chronic conditions in Year 1, but not in Year 2 (see Figure 1). These may represent coding gaps in Year 2 if Year 1 conditions were diagnosed and coded accurately because we would assume the chronic conditions to be also present in Year 2.

4. Recoded each of the flagged patients to add the chronic condition HCC to Year 2. This addition represents how the claims ideally should have been coded, to appropriately capture this clinical information. All other factors,

including demographics and other HCC categories in the risk adjustment model, were kept constant.

5. Reran the 2015 HHS-HCC risk adjustment model for Year 2, with the complete coding included.

6. Calculated the increase in risk score resulting from ensuring that patients with one of the nine chronic conditions in Year 1 also were coded with the same chronic condition in Year 2 (see Figure 2).

We can see from the table in Figure 1 that the observed coding persistency can vary greatly among the HCCs. Because we selected the HCCs to represent conditions that are chronic and therefore by definition present in both years, the persistency of less than 100% is indicative of not having claims that captured appropriate conditions in Year 2. Some are recoded from year to year quite often (such as HIV/AIDS and multiple sclerosis), and some are not (such as end-stage liver disease and fibrosis of lung and other lung disorders).

This analysis resulted in a 2.9% average risk score increase given the current benchmark diagnosis coding pattern. It is important to note that this increase is only from modeling the impact on these select nine HHS-HCC categories, whereas there do exist other conditions that are chronic and might similarly increase the average score if they were captured.

FIGURE 2: OVERALL RISK SCORE IMPACT IMPROVING CODING PERSISTENCY OF SELECT CHRONIC CONDITIONS

YEAR 2 ACTUAL AVERAGE RISK SCORE	1.501
YEAR 2 RECODED AVERAGE RISK SCORE	1.543
% INCREASE IN AVERAGE RISK SCORE	2.9%

FIGURE 1: PREVALENCE AND CODING PERSISTENCY BY SELECT HCC MEMBERS CONTINUOUSLY ENROLLED FOR TWO YEARS

	2015 HHS-HCC DESCRIPTION	ADULT MODEL GOLD RISK WEIGHT	ACTUAL PATIENT COUNT PER 10,000 IN YEAR 1	PERSISTENT PATIENT COUNT PER 10,000 IN YEAR 2 *	% PERSISTENCY
HHS_HCC001	HIV/AIDS	4.972	10.7	9.7	91%
HHS_HCC035	END-STAGE LIVER DISEASE	6.102	4.4	2.1	47%
HHS_HCC057	SYSTEMIC LUPUS ERYTHEMATOSUS AND OTHER AUTOIMMUNE DISORDERS	1.124	40.8	24.0	59%
HHS_HCC063	CLEFT LIP/CLEFT PALATE	1.978	2.5	1.8	70%
HHS_HCC088	MAJOR DEPRESSIVE AND BIPOLAR DISORDERS	1.698	427.3	283.9	66%
HHS_HCC102	AUTISTIC DISORDER	1.065	31.8	23.5	74%
HHS_HCC118	MULTIPLE SCLEROSIS	6.971	22.8	19.5	86%
HHS_HCC130	CONGESTIVE HEART FAILURE	3.648	62.8	38.5	61%
HHS_HCC162	FIBROSIS OF LUNG AND OTHER LUNG DISORDERS	2.657	16.6	8.9	54%

* Patients must be coded with the HCC in Year 1 to be included in Year 2 patient count.

FIGURE 3: USING RX CLAIMS TO INFER HCCS MEMBERS CONTINUOUSLY ENROLLED FOR TWO YEARS

	2015 HHS-HCC DESCRIPTION	ADULT MODEL GOLD RISK WEIGHT	ACTUAL PATIENT COUNT PER 10,000 IN YEAR 1	ACTUAL PATIENT COUNT PER 10,000 IN YEAR 1 + INFERRED FROM RX CLAIMS
HHS_HCC037	CHRONIC HEPATITIS (FOR HEP C ONLY)	1.228	22.1	22.3
HHS_HCC001	HIV/AIDS	4.972	10.7	13.2
HHS_HCC142	SPECIFIED HEART ARRHYTHMIAS	3.193	105.8	107.6
HHS_HCC048	INFLAMMATORY BOWEL DISEASE	2.640	61.7	77.7
G01	DIABETES	1.199	437.6	446.3
HHS_HCC118	MULTIPLE SCLEROSIS	6.971	22.8	23.7
HHS_HCC159	CYSTIC FIBROSIS	10.142	2.3	2.5
HHS_HCC130	CONGESTIVE HEART FAILURE	3.648	62.8	156.0

Using prescription drug claims to infer HHS-HCC categories

Using the same data set as above, we also identified potential coding accuracy issues caused by missing condition coding for prescription drug claims data. In other words, sometimes claims data includes a prescription drug, but not medical claims that correspond to what that drug is being taken for. For illustrative purposes, we used eight of the drug categories that HHS will include in the 2018 HHS-HCC risk adjustment model.[5] These eight categories correspond to a unique HCC or cluster of HCCs (in the case of diabetes) and can be used to infer the presence of an HCC in the absence of corresponding medical diagnoses in claims. The table in Figure 3 shows the prevalence rate for each of these eight categories in Year 1 (2014) before and after using prescription drug data. The table in Figure 4 shows the resulting increase in risk score if more accurate coding on medical claims were performed based on the information seen in the prescription drug data.

This analysis resulted in an overall risk score increase of 3.1%, all else being equal, if medical coding is tied to the prescription drug claims data for these eight categories.

FIGURE 4: OVERALL RISK SCORE IMPACT USING RX CLAIMS TO INFER HCCS

YEAR 1 ACTUAL AVERAGE RISK SCORE	1.384
YEAR 1 ADJUSTED AVERAGE RISK SCORE WITH RX DATA	1.426
% INCREASE IN AVERAGE RISK SCORE	3.1%

Conclusions

While actual results will vary by provider organization and patient population, our analysis shows that improvements in completeness and accuracy of condition information can have a meaningful impact on an organization's overall risk scores, and therefore on its financial performance in value-based contracting.

An analytic framework similar to what we have presented here can be used to identify areas for coding improvement, identify data issues, design targeted outreach and education to specific providers regarding their coding practices, and conduct retrospective evaluations of coding improvement efforts.

CONTACT

Jared Hirsch
jared.hirsch@milliman.com

Howard Kahn
howard.kahn@milliman.com

Rong Yi
rong.yi@milliman.com

FOR MORE ON MILLIMAN'S HEALTHCARE REFORM PERSPECTIVE:

Visit our reform library at **milliman.com/hcr**
Visit our blog at **healthcaretownhall.com**
Follow us at **twitter.com/millimanhealth**

---

5   For the 2018 benefit year, for commercial individual and small group markets subject to Patient Protection and Affordable Care Act (ACA) market rules, the Centers for Medicare and Medicaid Services (CMS) will apply a hybrid risk adjustment model that uses a combination of medical diagnoses and prescription drug data. For more information see the full HHS rule at **https://www.federalregister.gov/documents/2016/12/22/2016-30433/patient-protection-and-affordable-care-act-hhs-notice-of-benefit-and-payment-parameters-for-2018**.