Risk adjustment and the power of four

Ksenia Draaghtel, ASA, MAAA
Diane Laurent

For a long time, the healthcare industry has recognized the value of health status adjustments for predicting future healthcare costs, setting insurance premiums, and assuring fair comparisons and payments. Accuracy of risk adjustment systems and correct application of their results can have a significant consequence for the equity and efficiency of a healthcare system. Risk adjustment systems can be powerful instruments for assessing and adjusting health risks. Using risk adjustment to set fair and accurate premiums also enhances profitability, if the technology provides the necessary insight to improve pricing decisions. Unfortunately, many risk adjustment techniques currently in use have not evolved much over the past 15 years, and the adoption of risk adjustment by underwriters is not as prevalent as should be expected. Many of the older systems and methodologies are inadequate for pricing in the upcoming era of reforms.

State-of-the-art risk adjustment techniques, like Milliman's Advanced Risk Adjusters, can predict more than the traditional industry's 15% to 20%[1] of variance in individual health spending. An ideal risk adjuster would not only predict better, but would also provide more insightful scoring to give actuaries and underwriters information for optimal pricing.

With the advent of exchanges and other market reforms, carriers and providers would be wise to adopt more advanced risk adjustment techniques and ensure their proper implementation. This article highlights how advanced risk adjustment can be applied to set more accurate premiums; these advances can also help to shape other health reforms.

RELYING ON RISK SCORES BY SERVICE CATEGORY FOR RENEWAL UNDERWRITING SIGNIFICANTLY IMPROVES PROFITABILITY

Executive summary

Because of the improvement in predictive power that they offer, risk adjustment models have been extensively used for renewal underwriting in group health markets. The typical risk adjuster output offers a single risk score that represents an individual's overall health status risk. But segregating risk by service category better represents the differences in utilization and cost within each component, and is an important aspect in actuarial pricing. Inpatient, outpatient, physician, and pharmaceutical services possess different characteristics with respect to the utilization frequency, cost severity, speed of claim payment, and underlying trends. Hence, the ability to separate the risk by service category should allow for more accurate rate estimation, and is the subject of this case study.

To understand the financial impact of using a risk adjuster with risk scores by service category, two simulation case studies were conducted. In the first study, we analyzed profit for an incumbent carrier using one risk score where a competing, new carrier was using an age-sex manual rating approach. We compared this to the profit for an incumbent carrier using four service-category-specific scores when competing against the same new carrier. We found that, on average, the incumbent carrier using four scores earns a higher profit than the incumbent carrier using one score; the average difference was $6.49 per member per month (PMPM).

Our second study eliminates the new competing carrier and compares the profit position of the two incumbent carriers side by side (assuming they are competing with each other for the same business). In this comparison, the carrier using four scores earned an average profit $49.48 PMPM higher than the carrier using a single risk score.

These results suggest that there is a positive consequence to using a set of risk scores by service category, increasing the predictive power of a rating method, which may result in a significant impact on the bottom line.

Why split up the risk score?

As users of risk adjustment models have become more sophisticated, the use of risk scores has grown. The need for better models is more pronounced, as their applications have become more complex. Most traditional risk adjusters were developed during the late 1990s, when models were devised for general purposes: to profile providers or to estimate large-scale program budgets. These are the more common risk adjusters that output a single risk score to represent an individual's overall health risk. However, a single risk score does not provide a detailed description of the future expected risk, such as the source of risk by service category. The ability to separate the risk by service category allows for more accurate rate estimation. This paper will discuss these advantages and quantify the value to a health plan of using a risk adjuster model that provides risk scores by service category, as compared to a risk adjuster with a single risk score output.

Segregating healthcare utilization and costs by service category is an important aspect of sound actuarial pricing. Inpatient, outpatient, physician, and pharmaceutical services possess very different characteristics with respect to the utilization frequency, cost severity, speed of claim payment, and underlying trends. Hence, when projecting future claim costs, it is common to use

1 Society of Actuaries (2007). *A Comparative analysis of claims-based tools for health risk assessment.*

service-category-specific trend factors. Likewise, if a risk adjuster is used in making a projection of future claim costs, calculating the component of a risk score attributable to each service category before applying a service-category-specific trend will result in a better overall prediction.

Each service category has different predominant cost drivers and frequency of utilization. Hence the predictability varies by service category, which is illustrated in the table in Figure 1. This table displays metrics for the four service categories defined by the Milliman Advanced Risk Adjusters (MARA). For example, approximately 5% of the commercially insured population utilizes inpatient hospital services in a given year, which generally results in high claim costs. Consequently, because of the low-frequency nature of hospital inpatient claims, they are much harder to predict, reflected by a rather low percent of variance explained (R^2) by a risk adjuster model (6%). On the contrary, pharmaceutical services are sought much more frequently (by 75% of members) and generally result in a lower cost, and hence are much more predictable (R^2 value of 60%). Therefore, it is only logical that the amount of credibility that should be placed on experience versus risk score by service category should vary.

FIGURE 1: MEDICAL CLAIM STATISTICS BY SERVICE CATEGORY (INDIVIDUAL LEVEL MEASURES)

SERVICE CATEGORY	PROBABILITY OF CLAIM	PMPM COST	MARA R^2
INPATIENT FACILITY	5%	$183	6.0%
OUTPATIENT FACILITY	41%	$180	20.1%
PROFESSIONAL	89%	$245	20.2%
PHARMACEUTICAL	75%	$74	59.9%
TOTAL	95%	$682	28.8%

A low R^2 value reflects a low predictability of the future claim costs in a particular service category, lowering our confidence level in the risk score prediction as compared to the confidence level in the manual rate.

Two measures are commonly used in the industry to compare the accuracy of risk adjuster models—the R^2, which measures the percent of variance by the model, and the mean absolute percentage error (MAPE). Hence, a higher R^2 value and a lower MAPE are both indicative of a more accurate prediction. The table in Figure 2 presents R^2 and MAPE at the group level, for a sample population[2] for three rating methodologies—an age-sex manual rate, a traditional risk adjuster model with a single risk score, and a risk adjuster model with four risk scores by service category. The results in Figure 2 indicate that the risk adjuster model with four risk score outputs has the highest predictive accuracy among the three methods.

FIGURE 2: R^2 AND MAPE MEASURES BY RATING METHOD (GROUP LEVEL MEASURES)

RATING METHODOLOGY	MAPE	R^2
AGE-SEX RATING	34.6%	50.3%
SINGLE RISK SCORE	26.4%	73.4%
FOUR RISK SCORES BY SERVICE CATEGORY	25.7%	76.1%

Study results overview

Although the measures described above allow one to evaluate the relative accuracy of each method, they do not help to quantify the financial value of using the most accurate prediction. To understand the financial impact of using a risk adjuster with risk scores by service category, two simulation case studies were conducted. The first case study is a side-by-side comparison of two competitive scenarios where an incumbent carrier is competing with a new carrier for a group contract renewal:

1) Incumbent Carrier A uses a risk adjuster with a single risk score to develop the renewal rate, while the new carrier uses an age-sex manual rating.

2) Incumbent Carrier B uses a risk adjuster with four risk scores by four service categories (inpatient, outpatient, physician, and pharmacy), while the new carrier uses an age-sex manual rating.

In this case study, we found that, on average, Incumbent Carrier B's PMPM profit was higher by $6.49 PMPM than Incumbent Carrier A's profit. The table in Figure 3 presents a comparison of the resulting PMPM profit by group size for the two incumbent carriers.

FIGURE 3: CASE STUDY 1, PMPM PROFIT COMPARISON

GROUP SIZE	INCUMBENT B, AVG. PROFIT (PMPM)	INCUMBENT A, AVG. PROFIT (PMPM)	AVG. PROFIT DIFFERENTIAL (PMPM)
50-99	$74.27	$67.30	$6.98
100-249	$83.45	$77.88	$5.57
250-500	$91.22	$83.15	$8.07
AVERAGE	$83.62	$77.13	$6.49

In the second case study, we considered a scenario where two carriers were bidding for the same group of contracts, with the detailed past claim information available to both carriers. Carrier A used a risk adjuster with a single risk score, while Carrier B used a risk adjuster with four risk scores by service category. The average profit differential between the carriers was approximately $49.48 PMPM more for Carrier B than for Carrier A. The table in Figure 4 presents a comparison of the PMPM profit by group size for the two carriers.

2 Commercially insured population of 320,000 members belonging to 1,444 groups of sizes between 50 to 500 members.

Ksenia Draaghtel, Diane Laurent

FIGURE 4: CASE STUDY 2, PMPM PROFIT COMPARISON

GROUP SIZE	4-SCORE PMPM PROFIT	1-SCORE PMPM PROFIT	PROFIT DIFFERENTIAL
50-99	$42.03	($13.33)	$55.37
100-249	$45.36	$3.87	$41.49
250-500	$62.23	($0.14)	$62.37
AVERAGE	$49.06	($0.42)	$49.48

The results of the two studies lead to the conclusion that even a relatively small increase in the predictive power of a rating method may result in a significant impact on the profit bottom line and hence offer a competitive advantage.

Other industry studies

A similar simulation methodology has been used in past industry studies. In their 2003 study, "Applying Diagnostic-Based Predictive Models to Group Underwriting,"[3] Ellis and Kramer created a competitive simulation scenario to evaluate the economic value of a risk adjuster over traditional underwriting methods. In this scenario, there are two carriers, A and B, who are bidding for different group contracts. Carrier A uses a risk adjuster to price each group, while Carrier B uses age-sex and experience in its pricing. The carrier offering the lower price will win the contract. In the following year where the actual claim costs are known, the overall profits for Carrier A and Carrier B are calculated. This study found that for any group size (varying from 25 to 100), the risk adjuster renewal rating resulted in a higher PMPM profit than the traditional rating.

Assumptions and methodology

The study consisted of two scenarios. In the first scenario, an incumbent health plan has a group of contracts up for renewal and has to compete with a new carrier who will offer rates based on an age-sex basis. Incumbent Carrier A chooses to use a risk adjuster with a single overall risk score to estimate future claim costs, while Incumbent Carrier B chooses to use a risk adjuster with four risk scores to estimate future claim costs. The difference between the profitability of Incumbent Carriers A and B will illustrate the value of using a more advanced risk adjuster with detailed risk scores as compared to a single overall risk score rating. The diagram in Figure 5 illustrates this scenario.

In the second scenario, Carriers A and B are competing for the same block of business, assuming both have access to past detailed claim data in order to use risk adjuster tools. Carrier A uses a risk adjuster with a single risk score, while Carrier B uses a risk adjuster with four risk scores. The diagram in Figure 6 illustrates this scenario.

FIGURE 5: SINGLE OVERALL RISK SCORE VS. DETAILED MULTIPLE RISK SCORES

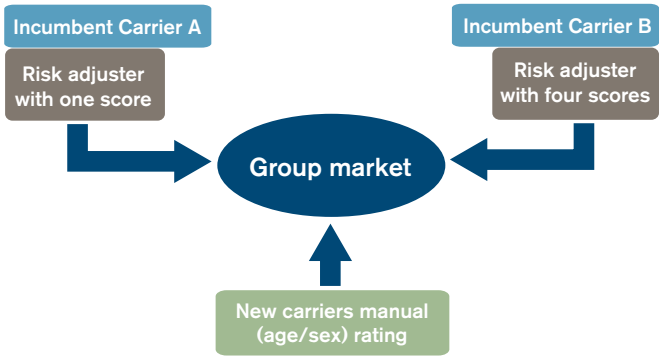


FIGURE 6: COMPETITIVE ENVIRONMENT SIMULATION

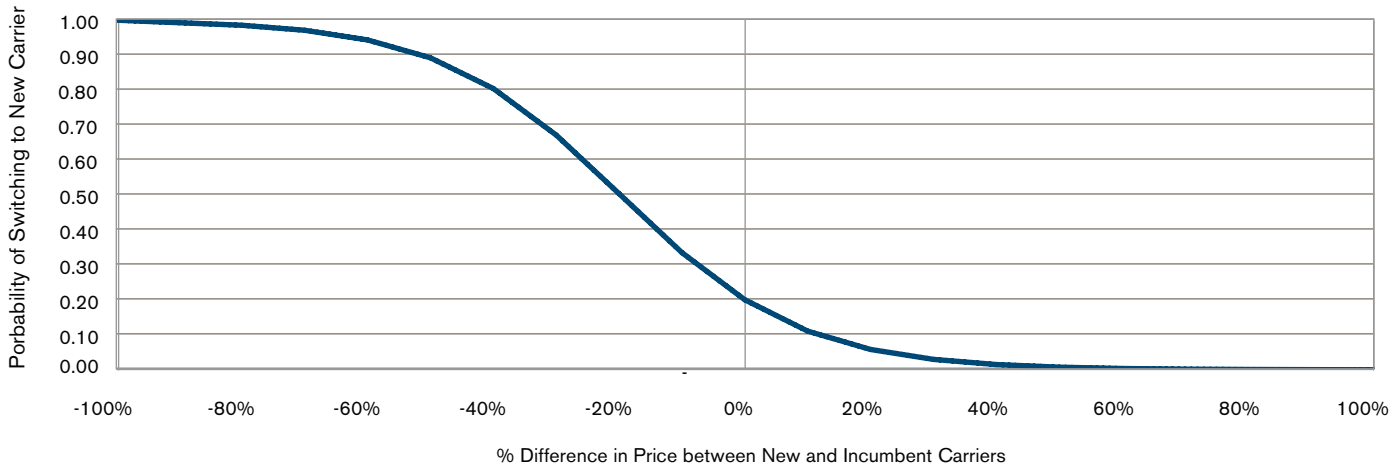
Bidder with lower price wins the contract

The study population consisted of 320,000 commercially insured members from the MarketScan claim database.[4] The population comprised 1,444 individual groups of 50 to 500 members, with two full years of eligibility in calendar years 2007 and 2008. The year 2007 was used as the assessment period and 2008 as the projection period; in other words, the claims in 2007 were used to project costs in 2008.

Even if a new carrier's offered rate is less than a group's current rate, this would not automatically mean that the group would drop existing coverage and switch to the new carrier. Factors other than price, such as brand loyalty or administrative cost of switching, play a role when a decision of coverage renewal is made, especially if the difference in price is not significant. We assumed that if there was no difference in price, 20% of groups will switch carriers, and we extrapolated a curve representing the probability of switching carriers given a difference in price between the new carrier's price and the incumbent carrier's price as a percentage of incumbent's price. Figure 7 presents the graph of the assumed curve. (See next page.)

3 Ellis, Randall, et al. (August 2003). Applying diagnosis-based predictive models to group underwriting. Health News Section, Issue 46.
4 MarketScan® Database, Thomson Reuters (Healthcare) Inc.

FIGURE 7: PROBABILITY OF SWITCHING CARRIERS VS. PRICE DIFFERENTIAL



For the new carrier's rating basis, we used age-sex factors based on the 2007 claim costs of the study population to estimate a group's average cost. The cost was then trended to the 2008 projection period using the actual historical trend of the population's claim costs from 2007 to 2008 of 11.9%.

For the incumbent carrier using a traditional risk adjuster with a single overall risk score (MARA can also produce a single risk score, and that is what was used for the traditional risk adjuster), the final rate was calculated as a weighted average of the 2007 experience and the risk-score-predicted cost with a risk score credibility factor of 87.5%. The credibility factor was estimated by minimizing the prediction error and was developed on a different population. Sensitivity testing of this assumption did not produce significant deviations in the overall result. The blended rate was then trended forward to the projection period using the actual 2007-to-2008 overall trend of 11.9%.

MARA was used to produce the four prospective risk scores by service category–inpatient (IP), outpatient (OP), physician (PHYS), and pharmaceutical services (Rx). Detailed medical claims in 2007 were classified into the four service categories to develop 2007 experience PMPMs claim costs by service category. Each service-category-specific risk score cost was blended with the experience PMPM by service category using the category-specific credibility weights. In addition, each service category was trended to 2008 using the category-specific cost trends, both of which are shown in Figure 8. The final projected cost was the sum of the projected cost for each service category.

FIGURE 8: COST TREND AND CREDIBILITY FACTORS BY SERVICE CATEGORY

	IP	OP	PHYS	RX	OVERALL
TREND	19.2%	13.0%	8.7%	8.2%	11.9%
CREDIBILITY FACTOR	100.0%	53.0%	61.5%	64.0%	87.5%

Using each of the three projection methods above, we measured the predictive accuracy of each rating method. The table in Figure 2 presents the MAPE and R^2 measures for each of the three methods described above.

Both risk adjuster methods outperformed the manual rating significantly, and the model with four risk scores by service category outperformed the model with a single overall risk score by 2.7 percentage points in R^2.

Finally, we assumed a 5% of premium profit margin and 15% of premium administrative expense load to create a gross rate for each group. Using the probability of switching and the actual 2008 projection year costs, the expected profit was calculated as the gross premium rate less the actual 2008 projection year costs, less the administrative expenses.

$$PROFIT_{CarrierX} = \sum_{All\ Groups} (\text{Gross Premium}_{\text{Carrier X, Group i}} - \text{2008 Actual Claim Costs}_{\text{Group i}} - \text{Expenses}_{\text{Carrier X, Group i}}) \times \text{Probability}_{\text{Group i selecting Carrier X}}$$

Study results and conclusions

The tables in Figures 9 and 1011 present the results of the first bidding scenario, where the two incumbent carriers competed with a new carrier. The risk adjuster model with four scores by service category resulted in a higher PMPM profit of $6.49 (or $301,554 in total monthly profit) compared to the risk adjuster model with a single risk score.

FIGURE 9: CASE STUDY 1, PMPM PROFIT COMPARISON

GROUP SIZE	INCUMBENT B, AVG. PROFIT (PMPM)	INCUMBENT A, AVG. PROFIT (PMPM)	AVG. PROFIT DIFFERENTIAL (PMPM)	AVG. PROFIT DIFFERENTIAL % OF A
50-99	$74.27	$67.30	$6.98	10.4%
100-249	$83.45	$77.88	$5.57	7.2%
250-500	$91.22	$83.15	$8.07	9.7%
AVERAGE	$83.62	$77.13	$6.49	8.4%

FIGURE 10: CASE STUDY 1, TOTAL MONTHLY PROFIT COMPARISON

GROUP SIZE	INCUMBENT B TOTAL MONTHLY PROFIT	INCUMBENT A TOTAL MONTHLY PROFIT	PROFIT DIFFERENTIAL
50-99	$2,298,611	$2,117,170	$181,441
100-249	$7,030,057	$6,666,191	$363,866
250-500	$3,649,347	$3,393,080	$256,267
AVERAGE	$5,215,064	$4,913,510	$301,554

The tables in Figures 11 and 12 present the results of the second bidding scenario, where the two incumbent carriers competed with each other using the same claim data for risk adjustment. The profit differential between the two carriers is dramatic, averaging $49.48 PMPM or $1,618,039 of total profit per month.

FIGURE 11: CASE STUDY 2, PMPM PROFIT COMPARISON

GROUP SIZE	4-SCORE PMPM PROFIT	1-SCORE PMPM PROFIT	PROFIT DIFFERENTIAL
50-99	$42.03	($13.33)	$55.37
100-249	$45.36	$3.87	$41.49
250-500	$62.23	($0.14)	$62.37
AVERAGE	$49.06	($0.42)	$49.48

FIGURE 12: CASE STUDY 2, TOTAL MONTHLY PROFIT COMPARISON

GROUP SIZE	4-SCORE TOTAL MONTHLY PROFIT	1-SCORE MONTHLY TOTAL PROFIT	PROFIT DIFFERENTIAL
50-99	$809,786	($288,432)	$1,098,218
100-249	$2,166,318	$244,758	$1,921,560
250-500	$1,475,782	($4,135)	$1,479,917
AVERAGE	$1,697,822	$79,784	$1,618,039

The results of the two case studies suggest that even a small improvement in the predictive power of a rating approach may lead to significant profit differential. The improvement in the predictive power of the risk adjuster model with the four risk scores by service category comes from greater refinement of the trend and credibility factors. There is significant variation in the trend and credibility factors by service category and this information is lost with a single risk score.

In reality, there are more than two carriers competing in the commercial small to mid-size group insurance market. In a highly competitive group market, the incumbent needs to leverage the past claim experience to the fullest extent in order to maintain competitive rates. A risk adjuster with four risk scores by service category enables a carrier to take advantage of the greater spectrum of information, delivers more complete information to the underwriter, and offers greater accuracy.

Ksenia Draaghtel, ASA, MAAA, is an associate actuary with the Denver office of Milliman. Contact her at *ksenia.draaghtel@milliman.com.*

Diane Laurent is the managing director of Milliman Advanced Risk Technologies in the firm's Minneapolis office. Contact her at *diane.laurent@milliman.com.*