JOURNAL OF HEALTH CARE FINANCE

**Fall 2019**

# Robust Risk Adjustment in Health Insurance

Tengjiao Xiao
Senior, Ernst & Young

Aurélie C. Thiele
Associate Professor, Southern Methodist University

## Abstract

**Objectives.** To create risk scores in a manner that incorporates uncertainty in the models used to redistribute payments to payers based on the risk of their enrollee population, to offset the cost of some payers' providing health insurance to relatively high-risk individuals.

**Study Design.** We propose a methodology based on worst-case optimization over an appropriate uncertainty set**.**

**Data Sources.** We use Federal Register data to create a realistic case study.

**Principal Findings.** Our approach to compute robust risk scores involves solving a series of linear problems and thus can be done using standard analytical tools such as Excel Solver.

**Conclusions.** There is an important need for the "robustification" of risk scores so that health payers can be properly compensated for the risk they take in providing insurance to relatively high-risk individuals. Our model presents one possible way to do so in a tractable manner.



**Keywords.** Risk adjustment, health insurance, robust risk scores.

## INTRODUCTION

Risk adjustment is "a statistical process used to identify and adjust for variation in patient outcomes that stem from differences in patient characteristics (or risk factors) across health care organizations." (Specifications Manual for National Hospital Quality Measures, [1]) Specifically, the risk adjustment program "provides payments to health insurance issuers that disproportionately attract higher-risk populations such as individuals with chronic conditions" and "transfers funds from plans with relatively lower risk enrollees to plans with relatively higher risk enrollees to protect against adverse selection." (Centers for Medicare and Medicaid Services, [2]) Without appropriate risk adjustment, comparing patient outcomes across organizations can be misleading because some top health organizations may consistently attract sicker patients, leading to lower health outcomes unrelated to the quality of care those patients receive. By accounting for existing risk factors, risk adjustment facilitates a fairer and more accurate inter-organizational comparison. The broad concepts and applications of risk adjustment are presented in Ellis [3].

Risk adjustment, as "an actuarial tool used to calibrate payments to health plans or other stakeholders based on the relative health of the at-risk populations" [4], can help remove the incentive for health plans to manipulate their offerings to deter the sick and attract the healthy [5]. Since insurers set premiums based on the riskiness of the people they enroll, adverse selection would also lead to higher premiums and government spending [6]. Glazer et al. [7] develop a statistical methodology to correct inefficient plan choice from adverse selection in health insurance markets, where enrollees sort between plans with fixed benefit offerings as a function of the plans' premiums. McWilliams et al. [8] show that the implementation of the Hierarchical Condition Categories (HCC) model is associated with reduced favorable selection in the Medicare Advantage program. Brown et al. [9] argue that risk adjustment can potentially increase the scope for payers' selecting individuals with costs below their capitation payment due to the increase in the variance of medical costs.

The official federal risk adjustment models, available in the Notice of Benefit and Payment Parameters of the Health and Human Services (HHS) [10], use fifteen weighted least squares regression models: platinum, gold, silver, bronze, and catastrophic for adult, child, and infant, respectively, to compute risk scores. The weight is the fraction of the year enrolled. Each HHS risk adjustment model predicts annual plan liability for an enrollee based on the person's age, gender, and diagnoses. The risk score of each enrollee equals the sum of all the risk weights associated with that patient, with the average risk score over the whole population being scaled to 1. The weighted average risk score of all enrollees in a particular health plan within a geographic rating area (the weights being again the fractions of the year enrolled) are then used as input to the payment transfer formula to determine an issuer's payment or charge for a plan, which is a baseline payment multiplied by the plan's enrollment-weighted average risk score [11].

The HHS risk adjustment model is a concurrent model, where diagnoses from a given period are used to predict cost in the same period. In contrast, a prospective model uses data from a prior period to predict costs in the current period or in the future. By design, both acute and chronic illnesses are emphasized in the concurrent model. In the prospective model, systematic factors, such as aging and chronic illnesses, outweigh acute and one-time conditions [12]. Acute and one-time events are averaged at the age/gender group level in the prospective model (Yi et al. [13]). The concurrent model is used by HHS because it is more robust to changes in enrollment than the prospective model ([13], [14]). In addition, prescription drugs are not included as a predictor in each HHS risk adjustment model. To evaluate model performance, $R^2$ and predictive ratios are examined, where the $R^2$ statistic calculates the percentage of individual variation explained by a model, and the predictive ratio is the ratio of the weighted-mean predicted plan liability to the weighted-mean actual plan liability for the model sample population [13].

Winkelman and Mehmud [15] use the Mean Absolute Prediction Error (MAPE) as an alternative to measure predictive accuracy. Glazer and McGuire [16] argue that, in order to address adverse selection and asymmetric information in managed care, risk adjustment should be viewed as a way to set prices for different individuals. Weiner et al. [17] quantify the impact of biased selection on health plans and evaluates mitigation attempts included in the Affordable Care Act.

Proper risk adjustment is thus very important for payers' long-term financial viability and for the competitiveness of the health insurance market. Risk adjustment has been used in the Medicare Advantage (MA) program, the Part D prescription drug program, many state Medicaid programs, the Commonwealth Care program in Massachusetts, and some employer-based plans [3]. Risk adjustment has also been implemented for the individual and small-group marketplaces. The main difference between the CMS-HCC model for Medicare and the HHS-HCC model for commercial insurance is that insurers get payments from CMS directly under Medicare, while payments are between insurers under commercial risk adjustment.

The weights for each risk factor can be obtained by linear regression, probit regression, or logistic regression, depending on the situation considered; however, estimates of regression coefficients are subject to error. Because risk adjustment in this context involves money transfers between health payers, and thus contributes to a payer's financial viability, it is critical to develop quantitative methods to incorporate ambiguity and uncertainty in the risk weights. The main contribution of this paper is to present a tractable methodology to create robust risk scores, which determine the amount of money to transfer between health payers and thus play an important role in their financial viability.

**Robustness in Risk Adjustment Models**

We first show the need for robustness on an example based on the Hospital Value-Based Purchasing (VBP) program, established by the Centers for Medicare and Medicaid Services (CMS) [18]. It aims at realigning hospitals' financial incentives by rewarding those that provide highest-quality care [19]. CMS funds the VBP adjustment scheme by withholding 1% of each hospital's Medicare payments, and re-distributing this pool of money to the hospitals based on the adjustment factors. Hospitals with the lowest adjustment factors receive little to no money back, and thus their 1% of Medicare payments will be lost to them and reassigned to better performing hospitals. Hospitals with the highest adjustment factors receive payments exceeding their initial 1% contribution to the pool. 1% might be ignored by bigger hospitals, but it can have a significant impact on smaller hospitals or hospitals in precarious financial health [20].

In our example, we investigate the variability between proxy and actual scores published by CMS [21] as follows. We first compute the rank of each hospital, based on the rank of its adjustment factor, with the hospital having the highest (best) adjustment factor receiving rank 1. We then merge the records under both the proxy and actual systems to compare proxy and actual ranks.

Ranks matter since the program is based on relative performance. The difference in rank is then computed as the proxy rank minus the actual rank, such that a positive difference represents a gain in ranks following the publication of the final (actual) factors. Figure 1 shows the differences in ranks from most negative to most positive on a representative year. Because the total number of hospitals is approximately 3,000, hospitals at the extreme left of the graph represent hospitals that had been expected to perform at the top based on proxy numbers and found themselves at the bottom when the actual numbers were published. Similarly, hospitals at the extreme right represent hospitals that had been deemed at the bottom based on the proxy factors and came out on top with the actual factors. 335 hospitals or 11.81% of the hospitals considered lost 1,000 spots or more and 250 hospitals or 8.81% gained 1,000 or more. The worst rank loss is a drop of 2,866 spots (from rank 21 to rank 2,887). The highest gain in rank (from rank 2,659 to rank 144) is an increase of 2,515. The wide fluctuation between the proxy factors and the actual ones has, to the best of our knowledge, not been discussed in the press or elsewhere, and suggests that there is a need to "robustify" factors.

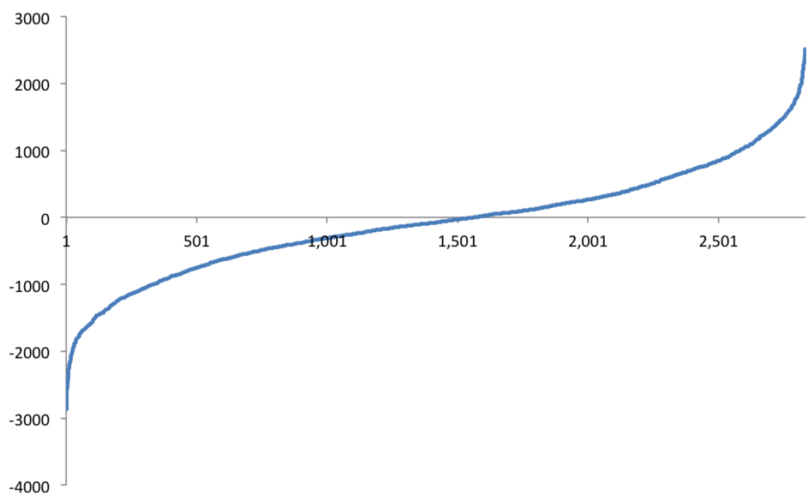


***Figure 1 Difference in Rank***

## Robust Risk Scoring

The traditional risk adjustment process, if the weights of the risk factors are known exactly, is:

1. Compute the risk score for each enrollee and scale it such that the average population risk score is one,
2. Compute the average risk score for each insurer (weighted by the fraction of the year each enrollee has been on the plan),
3. Determine the transfer payment as the difference between the insurer's cost (sum of patients' risk score times nominal cost) and his revenue (number of patients times capitated payment).

When the weights for the risk factors are not known precisely but estimates (for instance from a regression) and confidence intervals are available, we face the question of how this uncertainty should be incorporated so that payers receive a "fair" transfer payment. We will seek to minimize the *worst-case regret*. Here the worst-case regret is the greatest difference in absolute value between the estimated and actual risk scores computed over all payers and all possible weights for the risk factors within a predefined uncertainty set. It measures the worst-case difference in absolute value between the money transfer that should have taken place between payers if the true weights had been known and the transfer that actually did, based on the actual weights used to compute the risk scores. These weights are the decision variables of the problem.

We will use the following notation:

- $K$: the number of payers in the market,
- $S_k$: the set of enrollees of insurer k = 1,…,K,
- $J$: the set of conditions incorporated in risk scoring,
- $n_{jk}$: the number of enrollees of insurer k = 1,…,K who have condition j in J,
- $N_k$: the number of enrollees of plan k,
- $c_{ij}$ : a binary parameter equal to 1 when individual i has condition j,
- $w_j$ : the incremental risk weight for condition j in J (to be added to the risk score of individual i if $c_{ij} = 1$).

Insurer k's risk score before scaling is obtained by taking the average, over all enrollees, of the risk weights of the factors that affect the enrollee.

$$\frac{1}{N_k}\sum_{i \in S_k}\sum_{j \in J} w_j c_{ij} = \frac{1}{N_k}\sum_{j \in J} w_j n_{jk}$$

For convenience, we assume that all enrollees have been with the payer the whole year. Adapting the formulation to the case where some patients have joined the health plan during the year involves replacing the average over enrollees by a weighted average where the weights are the fractions of year for each patient. Risk scores are then scaled so that their population average is 1. Insurer k's average risk score after scaling becomes:

$$RS_k = \frac{\sum_{j \in J} w_j n_{jk}}{\sum_{l \in K}\sum_{j \in J} w_j n_{jl}} * \frac{\sum_{l \in K} N_l}{N_k}$$

We model the uncertain coefficients w's as belonging to a polyhedral set W. The set W can for instance be a box consisting of confidence intervals for each (independent) factor, or possibly include a budget-of-uncertainty constraint in the spirit of Bertsimas and Sim [22]. The problem we aim to solve in the decision variables v (the weights we want to give to each factor within the feasible set W) is then:

$$\min_{v \in W}\max_{k \in K}\max_{w \in W}\left|\frac{\sum_{j \in J} v_j n_{jk}}{\sum_{l \in K}\sum_{j \in J} v_j n_{jl}}\frac{\sum_{l \in K} N_l}{N_k} - \frac{\sum_{j \in J} w_j n_{jk}}{\sum_{l \in K}\sum_{j \in J} w_j n_{jl}}\frac{\sum_{l \in K} N_l}{N_k}\right| \quad (1)$$

Let's assume w.l.o.g that the polyhedral set W is represented as $\{w | l \le w \le u, Aw = b\}$. Further, let N be the $(n_{jk})$ matrix and e be the vector of all ones.

The key result of this section is the following theorem.

**Theorem 3.1 (Robust risk scoring).**

Problem (1) is equivalent to the linear problem:

$$\min_{x,y,Z} Z$$

$$s.t. Z \geq \frac{\sum_{l \in K} N_l}{N_k}\left(\sum_{j\in J} n_{jk}x_j - u_{-k}\right) \forall k$$

$$Z \geq \frac{\sum_{l \in K} N_l}{N_k}\left(-\sum_{j\in J} n_{jk}x_j + u_{+k}\right) \forall k$$

$$ly \leq x \leq uy, Ax = by, e'N'x = 1$$

where $u_{-k}$ and $u_{+k}$ are the respective optimal objectives values of the linear problems:

$$\min_{x,y} \sum_{j\in J} n_{jk}x_j$$

$$s.t.\ ly \leq x \leq uy, Ax = by, e'N'x = 1$$

and

$$\max_{x,y} \sum_{j\in J} n_{jk}x_j$$

$$s.t.\ ly \leq x \leq uy, Ax = by, e'N'x = 1$$

*Proof:* This follows from linearizing the piecewise linear term in (1) by introducing an auxiliary variable Z and linearizing the fractional terms by introducing the variables x and y, which are such that $x_j = \frac{v_j}{\sum_{l \in K}\sum_{j \in J} v_j n_{jl}}$ and $y = \frac{1}{\sum_{l \in K}\sum_{j \in J} v_j n_{jl}}$. These are classical transformations described in textbooks such as [23]. The reader is referred to Xiao [24] for details.

**Numerical Experiments and Discussions**

To test our approach, we generate a sample with 1,000,000 patients and 10 payers. The base payment is $2,000. The risk factors and nominal weights are taken from the Federal Register [10].

For illustrative purposes, the confidence interval of each risk weight is symmetric, centered at the nominal weight, and with a relative deviation from the mean selected randomly and up to 30% (i.e., the upper bound is at most 1.3 times the nominal weight.) The uncertainty set is a hypercube or "box" consisting of the range forecasts for each weight. Table 1 shows the nominal and robust weights as well as the lower and upper bounds of the weights used in the model. Table 2 compares nominal and robust risk scores for each insurer.

*Table 1 Nominal weights vs robust weights*

| Risk factor | Nominal weight | Deviation (%) | Lower bound | Upper bound | Robust weight |
|---|---|---|---|---|---|
| **Male, 21-24** | 0.258 | 24.44171 | 0.19494 | 0.32106 | 0.256263 |
| **Male, 25-29** | 0.278 | 27.17376 | 0.202457 | 0.353543 | 0.277949 |
| **Male, 30-34** | 0.338 | 3.809604 | 0.325124 | 0.350876 | 0.339511 |
| **Male, 35-39** | 0.413 | 27.40128 | 0.299833 | 0.526167 | 0.397926 |
| **Male, 40-44** | 0.487 | 18.97078 | 0.394612 | 0.579388 | 0.476095 |
| **Male, 45-49** | 0.581 | 2.926212 | 0.563999 | 0.598001 | 0.58552 |
| **Male, 50-54** | 0.737 | 8.354947 | 0.675424 | 0.798576 | 0.735302 |
| **Male, 55-59** | 0.863 | 16.40645 | 0.721412 | 1.004588 | 0.854907 |
| **Male, 60-64** | 1.028 | 28.72521 | 0.732705 | 1.323295 | 1.249658 |
| **Female, 21-24** | 0.433 | 28.94666 | 0.307661 | 0.558339 | 0.402888 |
| **Female, 25-29** | 0.548 | 4.728392 | 0.522088 | 0.573912 | 0.550033 |
| **Female, 30-34** | 0.656 | 29.11778 | 0.464987 | 0.847013 | 0.748113 |
| **Female, 35-39** | 0.76 | 28.71501 | 0.541766 | 0.978234 | 0.651303 |
| **Female, 40-44** | 0.839 | 14.56127 | 0.716831 | 0.961169 | 0.844894 |
| **Female, 45-49** | 0.878 | 24.00841 | 0.667206 | 1.088794 | 0.868385 |
| **Female, 50-54** | 1.013 | 4.25659 | 0.969881 | 1.056119 | 1.01657 |
| **Female, 55-59** | 1.054 | 12.65284 | 0.920639 | 1.187361 | 1.045615 |
| **Female, 60-64** | 1.156 | 27.47207 | 0.838423 | 1.473577 | 0.908063 |
| **Male, 2-4** | 0.283 | 23.76622 | 0.215742 | 0.350258 | 0.283589 |
| **Male, 5-9** | 0.196 | 28.78477 | 0.139582 | 0.252418 | 0.200356 |
| **Male, 10-14** | 0.246 | 19.67222 | 0.197606 | 0.294394 | 0.231851 |
| **Male, 15-20** | 0.336 | 1.07135 | 0.3324 | 0.3396 | 0.334237 |
| **Female, 2-4** | 0.233 | 25.47388 | 0.173646 | 0.292354 | 0.231625 |
| **Female, 5-9** | 0.165 | 28.0198 | 0.118767 | 0.211233 | 0.177605 |
| **Female, 10-14** | 0.223 | 20.36205 | 0.177593 | 0.268407 | 0.213277 |
| **Female, 15-20** | 0.379 | 22.7322 | 0.292845 | 0.465155 | 0.392902 |
| **Asthma** | 1.098 | 22.29397 | 0.853212 | 1.342788 | 1.12716 |
| **Acute Appendicitis** | 0.3 | 11.76681 | 0.2647 | 0.3353 | 0.319791 |
| **Diabetes** | 1.331 | 19.66434 | 1.069268 | 1.592732 | 1.554827 |
| **Congestive Heart Failure** | 3.79 | 5.135601 | 3.595361 | 3.984639 | 3.7253 |
| **HIV** | 5.485 | 21.18138 | 4.323201 | 6.646799 | 5.126647 |
| **Mental Illness** | 1.5 | 0.954985 | 1.485675 | 1.514325 | 1.499016 |

*Table 2 Nominal risk scores vs robust risk scores*

| | Insurer 1 | Insurer 2 | Insurer 3 | Insurer 4 | Insurer 5 |
|---|---|---|---|---|---|
| **Nominal Risk Score** | 0.9990379 | 0.999351 | 1.0005282 | 1.0008137 | 0.9995409 |
| **Robust Risk Score** | 0.9990487 | 0.9992846 | 1.0004502 | 1.0007771 | 0.9995645 |
| **Change in Risk Score** | 0.0011% | -0.0066% | -0.0078% | -0.0037% | 0.0024% |
| **Nominal Money Transfer** | -192413.36 | -129803.99 | 105639.84 | 162732.18 | -91822.65 |
| **Robust Money Transfer** | -190265.94 | -143077.69 | 90036.45 | 155424.44 | -87108.02 |
| **Change in Money Transfer** | -1.1160% | 10.2260% | -14.7704% | -4.4907% | -5.1345% |
| | **Insurer 6** | **Insurer 7** | **Insurer 8** | **Insurer 9** | **Insurer 10** |
| **Nominal Risk Score** | 1.0007512 | 1.0007304 | 1.0003757 | 1.0000186 | 0.9988525 |
| **Robust Risk Score** | 1.000775 | 1.0006677 | 1.0004926 | 1.0000597 | 0.9988801 |
| **Change in Risk Score** | 0.0024% | -0.0063% | 0.0117% | 0.0041% | 0.0028% |
| **Nominal Money Transfer** | 150238.92 | 146073.36 | 75135.25 | 3714.9 | -229494.45 |
| **Robust Money Transfer** | 154993.01 | 133534.91 | 98516.48 | 11931.7 | -223985.34 |
| **Change in Money Transfer** | 3.1644% | -8.5837% | 31.1189% | 221.1848% | -2.4005% |

We can see from Table 2 that although the percentage changes in risk scores are small, the changes in actual money transfers are significant. The reason is that the relative change in risk score is calculated as (RS-RSN)/RSN, while the relative change in actual money transfer is calculated as [(RS-1)*N*C-(RSN-1)*N*C]/(RSN-1)*N*C , or equivalently (RS-RSN)/(RSN-1): the numerator stays the same but the denominator does not and this can create significant changes because the risk scores are close to 1 to begin with. In the example above, 4 out of 10 payers observe a relative change in actual money transfer higher than 10% and Insurer 8 sees a 72% increase in payment.

## CONCLUSIONS

In this paper, we have investigated how to mitigate the impact of uncertainty on the estimates of risk factors for adjustment models in healthcare. Risk adjustment involves scoring the enrollee population of each payer to account for the population's health status and deciding transfer payments between payers, so that health plans are rewarded for appropriate care but not for enrolling healthier patients than their competitors. We provided an example related to hospital ranking to demonstrate the need for robustness. We then presented an approach to compute robust risk scores. Our methodology involves solving a series of linear problems and thus is easy to implement using standard analytical software.

## REFERENCES


[1] Staff, ORYX risk adjustment guide, Tech. rep., Joint Commission (2014).
[2] Staff, Risk Adjustment Overview, Centers for Medicare and Medicaid Services (2012).
[3] R. Ellis, Risk adjustment in health care markets: concepts and applications, Wiley (2008).
[4] Staff, Risk assessment and risk adjustment, American Academy of Actuaries Issue Brief.
[5] D. M. Cutler, R. J. Zeckhauser, Adverse selection in health insurance, Frontiers in Health Policy Research, volume 1, MIT (January 1998).
[6] L. J. Blumberg, J. Holahan, The affordable care act can survive low enrollment and adverse selection in the first year, Timely analysis of immediate health policy issues, Robert Wood Johnson Foundation (December 2013).
[7] J. Glazer, T. G. McGuire, J. Shi, Risk adjustment of health plan payments to correct inefficient plan choice from adverse selection, Working Paper 19998, National Bureau of Economic Research (March 2014).
[8] J. M. McWilliams, J. Hsu, J. Newhouse, New risk-adjustment system was associated with reduced favorable selection in Medicare Advantage, Health Affairs 31 (12) (2012) 2630-2640.
[9] J. Brown, M. Duggan, I. Kuziemko, W. Woolston, Will risk adjustment decrease health care costs? New evidence from the Medicare Advantage program, Tech. rep., National Bureau of Economic Research (2011).
[10] Staff, Department of Health and Human Services 45 CFR parts 153, 155, 156, 157 and 158, Federal Register 77 (236).
[11] Staff, Issues for risk adjustment in Medicare Advantage, Tech. rep., Medpac (2012).
[12] Staff, Risk adjustment implementation issues, Tech. rep., Centers for Medicare and Medicaid Services (2011).
[13] R. Yi, J. Shreve, W. Bluhm, Risk adjustment and its applications in global payments to providers, Milliman client report, Milliman, Inc. (2011).
[14] R. Winkelman, R. Damler, Risk adjustment in state Medicaid programs, Health Watch.
[15] R. Winkelman, S. Mehmud, A comparative analysis of claims-based tools for health risk assessment, Tech. rep., Society of Actuaries (April 2007).
[16] J. Glazer, T. McGuire, Optimal risk adjustment in markets with adverse selection: An application to managed care, The American Economic Review 90 (4) (2000) 1055-1071.
[17] J. Weiner, E. Trish, C. Abrams, K. Lemke, Adjusting for risk selection in state health insurance exchanges will be critically important and feasible, but not easy., Health Affairs 31 (2) (2012) 306-315.
[18] Staff, Value-based purchasing: a strategic overview for health care industry stakeholders, Tech. rep., Deloitte Center for Health Solutions (2011).
[19] J. Rau, Methodology: How value-based purchasing payments are calculated, Tech. rep., Kaiser Health News (2013).
[20] Staff, Profitability and revenue growth drop in US not-for-profit hospital preliminary medians, Tech. rep., Moody's Investors Service (2014).
[21] Staff, Details for title: FY 2013 final rule tables, Tech. rep., Centers for Medicare and Medicaid Services, http://go.cms.gov/1sCr99T (2013).
[22] D. Bertsimas, M. Sim, The price of robustness, Operations Research 52 (1) (2004) 35-53.
[23] D. Bertsimas, J. Tsitsiklis. Introduction to Linear Optimization. Athena Scientific. Belmont, MA (1997)
[24] T. Xiao. Robust Healthcare Financing Systems. PhD Dissertation, Lehigh University, Bethlehem, PA. (2015)