March 31, 2016, HHS-Operated Risk Adjustment Methodology Meeting

Discussion Paper

March 24, 2016

Centers for Medicare & Medicaid Services

Center for Consumer Information & Insurance Oversight

Contents

List of Tables

EXECUTIVE SUMMARY

Starting with coverage beginning in 2014, individuals and small businesses are able to purchase private health insurance through reformed markets that encourage competition on quality and value. The Affordable Care Act established a permanent risk adjustment program to minimize the negative effects of adverse selection and help level the playing field between insurance companies, thereby fostering a stable, vibrant market in which issuers are rewarded for providing high-quality, affordable coverage, not for offering plans designed to attract the healthy and avoid the sick. The program applies to non-grandfathered health plans in the individual and small group markets, inside and outside the State-based Marketplaces and the Federally-facilitated Marketplaces (the Marketplaces). The risk adjustment program is intended to achieve this goal by mitigating the effect of risk selection on premiums by transferring premium revenue from plans with below-average actuarial risk to plans with above-average actuarial risk. Such a transfer mechanism is an essential component of the insurance market reforms implemented by the Affordable Care Act. These market reforms include:

- **Guaranteed issue/renewal.** All non-grandfathered insurance coverage offered by health insurance issuers must be offered on a guaranteed issue basis. Health insurance issuers may not refuse to issue or renew coverage to any individual on the basis of their health status or prior use of health services.

- **Adjusted community rating**. A health insurance issuer may not charge an individual more for non-grandfathered individual or small group market coverage based on that individual's health status or prior use of health services. A health plan may vary the premium it charges a subscriber based only on four factors: 1) family structure/size, 2) geographic rating area, 3) age, within a factor of 3:1 among adults, and 4) tobacco use within a factor of 1.5:1. Each health plan will, consistent with State regulations, determine the premium it charges for each of its products.

- **Single risk pool.** The single risk pool requirement directs an issuer to develop its market index rates for non-grandfathered insurance plans in the individual and small group markets based on the pooled essential health benefits claims experience of all of its enrollees in all non-grandfathered health plans in the applicable market in a State. After setting the index rate, the issuer must make a market-wide adjustment based on the expected aggregate payments and charges under the risk adjustment and reinsurance programs for all plans covered by those programs in a market within a State and fees. The premium rate for any given plan cannot vary from the resulting index rate except for plan-specific adjustments for: the actuarial value and cost-sharing design of the plan; the plan's provider network and delivery system characteristics and utilization management practices; plan benefits provided in addition to the essential health benefits; administrative costs; and with respect to catastrophic plans, the expected impact of the specific eligibility categories for those plans.

These three provisions mean that a health plan that enrolls individuals in poorer health would not be able to charge higher premiums than an otherwise identical health plan that enrolls healthier individuals. In the absence of a risk adjustment mechanism, a health plan would gain a

competitive advantage if it enrolled the healthy and avoided the sick. This could create an incentive for issuers to avoid offering plan designs that are particularly valuable to sicker individuals, thereby reducing the variety and quality of the coverage consumers have to choose from. The goal of the Affordable Care Act market reforms, and the goal of its risk adjustment program, is to create a stable market in which health plan premiums will reflect the value of the coverage offered, including product features such as effective care delivery, and not risk selection.

The risk adjustment program is intended to provide increased payments to health insurance issuers that attract higher-risk populations (such as those with chronic conditions) and reduce the incentives for issuers to avoid higher-risk enrollees. Under the risk adjustment program, funds are transferred from issuers with lower-risk enrollees to issuers with higher-risk enrollees in a budget-neutral manner, thereby reducing influence of risk selection on the premiums that plans charge. The Affordable Care Act authorizes the Department of Health and Human Services (HHS) to utilize criteria and methods similar to those utilized under Medicare Parts C or D to implement risk adjustment. The HHS risk adjustment methodology developed by the Centers for Medicare & Medicaid Services (CMS) and its contractor, RTI International, is based on the premise that premiums should reflect the differences in plan benefits, quality, and efficiency, and not the health status of the enrolled population. The HHS risk adjustment methodology includes the risk adjustment model and the payment transfer formula.

The HHS risk adjustment model uses an individual's demographic data and diagnoses to determine a risk score, which is a relative measure of how costly that individual is anticipated to be to the plan (i.e., a relative measure of the individual's actuarial risk to the plan). Risk adjustment modeling determines the base actuarial risk based on predicted costs for a plan's enrollees. The payment transfer formula applies that actuarial risk in a manner that accounts for many other factors in an issuer's enrollee population in the market and State to determine the issuer's transfer amount.[1] The payment transfer formula is based on the difference between two plan premium estimates: 1) the predicted premium with risk selection, and 2) the predicted premium without risk selection. Transfers are intended to bridge the gap between these two premium estimates. The goal of risk adjustment transfers is to calculate balanced transfers that account for health risk differences while preserving permissible premium differences.

HHS has actively sought comment and received feedback on the risk adjustment methodology from the beginning of the program's development. After the implementation of the risk adjustment program for the 2014 benefit year, we received formal and informal feedback, including recommendations that the model should include prescription drugs as predictors of diagnoses or severity of illness; the model should better account for partial year enrollment; and the model should be recalibrated on the individual and small group populations rather than a separate commercial dataset. We have also received comments about how the methodology affects smaller or more efficient issuers. We explore these issues below.

This paper aims to provide the public with a cohesive summary of the risk adjustment methodology, including detailed explanations of the risk adjustment models and the payment transfer formula, as well as the updates to the model we have made since the initial 2014

[1] These other factors include plan allowable premium rating, actuarial value, induced demand, geographic costs, market share, and the Statewide average premium.

calibration. We also explore potential modifications to the risk adjustment methodology for the 2018 benefit year and beyond. We continue to seek feedback from the public on these issues, as well as other modifications to our risk adjustment models and the payment transfer formula that CMS should consider, including with respect to how these various changes will interact. HHS is holding a public meeting on Thursday, March 31, 2016, to discuss these potential proposals. We are requesting comments on the proposals discussed in this paper through hhshccraops@cms.hhs.gov or REGTAP at https://www.REGTAP.info,[2] by April 22, 2016. Comments sent in direct response to this paper will inform future rulemaking and continued evaluation of the methodology. This paper was drafted as part of our ongoing engagement with States, issuers, consumers, and the public on this important program. The considerations described in this paper are intended for discussion, and do not represent determinations of the policies that CMS will or will not propose or pursue.

In Chapters 1 through 3 of this paper, we describe the key program goals that informed development of the HHS risk adjustment models, summarize the HHS hierarchical condition categories (HCC) diagnostic classification, provide an overview of the data and methods that were used to develop the risk adjustment model for each age group (adult, child, and infant) and metal level (platinum, gold, silver, bronze as well as catastrophic), and discuss the updates to the risk adjustment model over the years since the model's first development and implementation for the 2014 benefit year and through the 2015, 2016 and 2017 benefit years.

Chapter 4 discusses the proposals we are considering to improve the model's ability to predict risk. In Chapter 4, we discuss the following topics:

1) Whether and how to account for partial year enrollment, whether through separate risk adjustment models based on enrollment duration or using interaction factors developed by type of condition;

2) Whether and how to develop a prescription drug model; specifically, a "hybrid" drug-diagnosis model. We describe the benefits and concerns regarding adding prescription drug utilization to a diagnosis-based risk adjustment model, describe the empirical framework for inclusion and the drug classification and aggregation systems, and identify criteria for selecting the drug-diagnosis pairs for inclusion in the model;

3) Whether and how to pool high risk enrollees in HHS risk adjustment;

4) An evaluation of concurrent and prospective risk adjustment models;

5) Recalibration of the 2018 risk adjustment model using the most recent data; and

6) Evaluating the current distributed data environment and data collection (EDGE servers), including the benefits of basing risk adjustment models on individual and

[2] To submit comments, an individual must register as a REGTAP user. Once registered, you can access the 'Submit an Inquiry' feature on the REGTAP Dashboard. In the 'Submit an Inquiry' feature, please enter your comment, select the 'Content' button, and choose 'HHS-Operated Risk Adjustment Methodology Meeting (03/31/2016)' as the event title. If you have any attachments, you can upload an attachment within this feature. Please be sure to enter only one (1) comment per submission.

small group Affordable Care Act-compliant data, additional variables that could be collected to allow recalibration of the HHS risk adjustment models, and the enrollee-level data collection process, burden, timing, privacy and security, and data use issues.

In Chapter 5, we present the payment transfer formula. Key issues for HHS risk adjustment are that risk adjustment transfers must be balanced within a risk pool within a market within a State (i.e., must sum to zero across all plans), that health status is not an allowable rating factor, and that plan premiums for adults must not vary by age by more than a 3 to 1 ratio. We first describe how the plan risk score is combined with factors for plan allowable premium rating, actuarial value, induced demand, geographic costs, and the Statewide average premium in a formula that calculates transfers among plans. We then show how each plan factor is determined, as well as how the factors relate to each other in the transfer formula.

We also address the feedback we have received regarding risk adjustment transfers and the transfer formula, including the inclusion of administrative expenses in the Statewide average premium, potential modifications to the transfer formula, and our initial findings from the 2014 benefit year. We describe the outcomes of risk adjustment in the individual and small group markets for the 2014 benefit year. We provide an overview of 2014 risk adjustment transfers at the issuer and plan levels as a percent of premiums. We describe transfers by plan metal level and plan liability risk score. Finally, we examine risk adjustment transfers incorporating reinsurance payments in the individual market.

CHAPTER 1: HHS RISK ADJUSTMENT METHODOLOGY OVERVIEW

1.1 Background

Section 1343 of the Affordable Care Act provides for a permanent risk adjustment program for all non-grandfathered plans in the individual and small group market both inside and outside of the Marketplaces. The Affordable Care Act directs the Secretary, in consultation with the States, to establish criteria and methods to be used in determining the actuarial risk of plans within a State. States electing to operate a risk adjustment program, or HHS on behalf of States not electing to operate a risk adjustment program, will assess charges to plans that experience lower than average actuarial risk and use the collected charges to make payments to plans that have higher than average actuarial risk.

Without risk adjustment, plans that enroll a higher proportion of high-risk enrollees would need to charge a higher average premium (across all of their enrollees) to be financially viable. Enrollees in health insurance plans differ in their expected cost, or risk, because of differences in their health status. The intent of risk adjustment is to allow a plan enrolling a higher proportion of high-risk enrollees to charge the same average premium (other factors being equal) as a plan enrolling a higher proportion of low-risk enrollees, shifting the focus of plan competition to plan benefits, quality, efficiency, and value.

Risk adjustment is widely recognized as a critical component of competitive health insurance markets. The Medicare Advantage program, through which private plans provide health insurance to Medicare beneficiaries, utilizes a risk adjustment mechanism (Pope et al., 2004), as does the Medicare Part D program, which provides prescription drug insurance through private plans to Medicare beneficiaries (Kautter et al., 2012). Many State Medicaid programs utilize risk adjustment as well (Winkelman and Damler, 2008). In addition, several countries, including the Netherlands, Switzerland, Germany, Ireland, Australia, and South Africa, have introduced a risk adjustment mechanism as part of their regulated private health insurance markets (Armstrong et al., 2010; Schokkaert, 2006). Prior to the enactment of the Affordable Care Act, risk adjustment was not commonly used in United States private health insurance markets, though risk adjustment was used in Massachusetts for the Commonwealth Care program beginning in 2009.[3]

The HHS risk adjustment methodology includes the risk adjustment models and the payment transfer formula, or "transfer formula," and was established in the HHS Notice of Benefit and Payment Parameters for 2014 (2014 Payment Notice) (78 Federal Register 15410).[4] The risk adjustment model uses an individual's demographics and diagnoses to determine a risk score, which is a relative measure of how costly an individual is anticipated to be to the plan (i.e., a relative measure of the individual's actuarial risk to the plan). The transfer formula averages all individual risk scores in risk adjustment covered plans, and uses the plan average risk scores,

[3] Commonwealth of Massachusetts Notice of Benefit and Payment Parameters 2014. "Risk Adjustment Methodology and Operations." April 2013. https://betterhealthconnector.com/wp-content/uploads/reports-and-publications/Risk_Adjustment/MANoticeofBenefitPaymentParameters.pdf

[4] The risk adjustment methodology also includes the *data collection approach* and the *schedule for the risk adjustment program*. Please note this is the methodology used by HHS when operating risk adjustment on behalf of a State. For benefit years 2014 through 2016, Massachusetts operates its own risk adjustment methodology.

combined with other factors,[5] to calculate the money transferred between plans. The risk adjustment payment transfer formula is based on the difference between two plan premium estimates: 1) premium with risk selection, and 2) premium without risk selection. Transfers are intended to bridge the gap between these two premium estimates. Conceptually, the goal of risk adjustment transfers is to calculate balanced transfers that account for health risk differences while preserving permissible premium differences. Here, we discuss the policy goals and motivations behind the 2014 HHS risk adjustment methodology.

1.2 HHS Risk Adjustment Development: Goal and Issues

The key program goal of the HHS risk adjustment methodology is to compensate health insurance plans for differences in enrollee health mix so that plan premiums reflect differences in scope of coverage and other plan factors, but not differences in health status of enrollees. The methodology addresses three issues specific to HHS risk adjustment for State individual and small group markets: 1) new population; 2) market factors; and 3) balanced transfers within State/market/risk pool.

1.2.1 New Population

When the methodology was developed, the HHS risk adjustment population was unknown. It was assumed the newly-constituted population would be defined by who enrolled in the State individual and small group markets, inside and outside the Marketplaces, beginning in 2014. The new population would include not only those who previously had private (or public) health insurance coverage, but also individuals who were previously uninsured. As a new population, medical claims data for the risk adjustment population was not available for use in calibrating a risk adjustment model. A proxy source of data was needed to calibrate the risk adjustment model. Medicare data were not appropriate because the HHS risk adjustment population is largely under age 65 and has a large proportion of employed enrollees. Instead, data from employer-sponsored insurance or Medicaid were the most likely sources of calibration data. Another consideration was that even after risk adjustment was implemented, some enrollees subject to HHS risk adjustment would have partial year enrollment periods even in the current year if they transition to or from Medicaid or large-employer-based insurance.

1.2.2 Market Factors – Actuarial Value and Permissible Rating

Different Plan Actuarial Value Levels versus a Standard Benefit Level. The Affordable Care Act established four tiers of plan actuarial value, or "metal levels," plus catastrophic plans, which are risk adjusted in a separate risk pool. The metal levels are platinum, gold, silver, and bronze, which correspond, respectively, to plans that are estimated to pay 90, 80, 70, and 60 percent of the medical expenditures of a standard population.

The presence in the market of plans with different actuarial values posed a challenge for the risk adjustment methodology – how to preserve premium differences that reflect differences in generosity of plan coverage. Risk adjustment transfers should counteract the effects of risk selection, but not differences attributable to the fact that identical individuals enrolled in a higher actuarial value plans will pay lower cost sharing or use more services because of that lower cost

[5] These other factors include plan allowable premium rating, actuarial value, induced demand, geographic costs, market share, and the Statewide average premium.

sharing. Other things being equal, an individual should pay a higher premium for a platinum plan than a bronze plan to reflect the reduced cost sharing the individual pays when enrolled in the platinum plan. However, an individual should not pay more to enroll in a platinum plan because it has sicker enrollees on average than the bronze plan.

Allowed Rating Factors versus Uniform Premiums. The Affordable Care Act allows individual and small group plans to rate premiums based on four factors: family size, geographic rating area, age, and tobacco use. The age variation in premiums is constrained to 3:1 for enrollees aged 21 and older, and the variation based on tobacco use is constrained to 1.5:1. In the presence of age rating variation, if a plan obtains higher revenues by charging its older enrollees more, that revenue difference should be accounted for in its risk adjustment transfers. Age is a proxy for the prediction of medical expenditures and is typically included in risk adjustment models. Families with three or fewer children are calculated as the sum of individual rates for each individual within the family, based on each person's age, tobacco use, and geographic rating area. For families with more than three children, the family premium would be built from individual premiums of the parents plus the three oldest children, with additional children not reflected in the family premium.

Geographic rating area is the final source of allowed rating variation. Individual and small group markets are established within States, but States may elect to define multiple intra-State rating areas across which plans can vary premiums.[6] We needed to consider how risk adjustment transfers should differ based on different premium and cost levels across rating areas. More generally, we evaluated how a methodology could be established that is flexible enough to potentially be applied to all 50 States, with their different cost levels.

1.2.3 Balanced Risk Adjustment Transfers among Plans versus Risk-Adjusted Payment to Plans

Determining how to calculate balanced risk adjustment transfers among plans while preserving permissible premium differences was a central task we faced in developing the HHS risk adjustment methodology. In the Affordable Care Act-defined individual and small group markets, risk adjustment determines transfers among health insurance plans. Lower risk plans are charged to fund payments to higher risk plans. The payments and charges are balanced (i.e., the transfers sum to zero). HHS risk adjustment reallocates aggregate premium revenue among plans, whether premiums are paid by individual enrollees or the government through income-based subsidies.

1.3 HHS Risk Adjustment Development: Approach

The risk adjustment methodology includes a risk adjustment model and a payment transfer formula that together address the key goals and issues discussed above. The risk adjustment model estimates differences in health risks taking into account the new population and generosity of coverage (actuarial value level). The payment transfer formula includes the plan risk score and other cost factors, relative to the market average, to calculate balanced transfers that are intended to account for health risk differences while preserving permissible premium differences.

---

[6] State rating areas are subject to approval by HHS.

1.3.1 HHS Risk Adjustment Model

The HHS risk adjustment model uses an individual's demographics and diagnoses to determine a risk score, which is a relative measure of how costly that individual is anticipated to be to the plan (i.e., a relative measure of the individual's actuarial risk to the plan). The model was developed by estimating how demographics (age and sex) and health diagnoses relate to health expenditures. Below we describe several features of the model that address the new population and plan actuarial value differences described above.

Employer-Sponsored versus Medicaid Data to Calibrate a Risk Adjustment Model. At the time the HHS risk adjustment model was developed, projections of the characteristics of the long-run (2019) individual market population (both inside and outside the Marketplaces) had been made in comparison to the surveyed employer-sponsored insurance and Medicaid enrollees (Trish, Damico, Claxton, et al., 2011). Although many projected characteristics of the individual market enrollees were similar to those of both enrollees in employer-sponsored insurance and Medicaid enrollees, on average they tended to be closer to enrollees in employer-sponsored insurance. In addition, many of the commercial issuers that enrolled the employer-sponsored population would enroll the HHS risk adjustment population, and their provider payment rates were expected to be similar. In contrast, Medicaid fee-for-service payment rates are set by States, and are often well below commercial issuer payment levels. For these reasons, we focused on claims data from employer-sponsored insurance to calibrate the HHS risk adjustment models. The specific employer-sponsored insurance claims dataset we chose, Truven MarketScan® Commercial Claims and Encounter data, is discussed below in Chapter 3.[7]

Prospective versus Concurrent Model. Risk adjustment models can only utilize available information to predict expenditures. Most risk adjustment models used for payment are "prospective," meaning they use prior year information to predict current year medical expenditures. For example, the Medicare Advantage, Medicare Part D, and the Netherlands' risk adjustment models are prospective. Prospective models tend to emphasize the impact of ongoing chronic conditions on costs (as opposed to random current year costs that can be pooled as "insurance risk"). No previous year information on health status existed for the first year of the Affordable Care Act-established individual and small group markets in 2014. Additionally, unlike with Medicare, enrollees move in and out of enrollment in the individual and small group markets and move across issuers. A prospective model was, therefore, infeasible for the first year of the Affordable Care Act. Additionally, given the time required to accumulate and analyze data and pre-announce the model and our de-centralized data collection process, it continues to be infeasible for the foreseeable future. Because a prospective model cannot easily reflect enrollees' movement between markets and across issuers, we believe a concurrent model is more appropriate for the individual and small group markets.

Concurrent models tend to emphasize the prediction of costs associated with current year acute health events. A considerable amount of the costs of chronic conditions are associated with acute exacerbations, which a concurrent model will better capture. Concurrent models can also capture the very high costs of conditions such as organ transplants, metastatic cancer, and low-birthweight babies that reduce or eliminate the disincentive for plans to contract with providers that treat these conditions

[7] Note that IBM recently purchased Truven.

In developing the concurrent model, we attempted to focus on conditions associated with systematic selection risk of enrollees or providers, and to de-emphasize conditions such as injuries that are probably not a focus of plan selection behavior. Further, because concurrent risk adjustment explains more of the variation in current (acute) costs, it reduces unsystematic risk, which may benefit small health plans that do not have enough enrollees to diversify away unsystematic risk. Finally, we included partial year enrollees in the sample to calibrate the risk adjustment model because, with a concurrent risk adjustment model, enrollees' diagnoses will match their utilization for any period of enrollment. This means that newborns and decedents, some of whom are typically among the highest-cost enrollees, are included. For these reasons, all enrollees (with at least one month of enrollment) were included.

Revised Clinical Classification and Subpopulation Models. The HHS risk adjustment model predicts expenditures using only enrollees' age, sex, and diagnoses. A diagnosis is a key clinical factor that drives medical treatment decisions and costs, and is widely used in risk adjustment models (American Academy of Actuaries, 2010). Conceptually, a diagnosis is distinct from treatment or utilization, and we believe basing risk adjustment on diagnoses is neutral with respect to treatment modality and utilization. The heart of the risk adjustment model is the clinical classification system that organizes the thousands of International Classification of Diseases (ICD) diagnosis codes into a coherent system of diagnostic categories.

The Centers for Medicare & Medicaid Services Hierarchical Condition Categories (CMS-HCC) clinical classification (Pope et al., 2004) was the starting point for the HHS risk adjustment diagnostic clinical classifications. [8] The base CMS-HCC model used to develop the HHS-HCC model included the entire clinical classification structure of the CMS-HCC model, including not only the 87 HCCs used in Medicare payment, but also 114 HCCs that are not used to determine Medicare payment but are part of the underlying model. By including all payment and non-payment HCCs in the analysis, we made an independent determination regarding the structuring of the HCCs as well as which HCCs would be included for payment in the HHS-HCC model, building a new model that was specifically designed for the Affordable Care Act population. The CMS-HCCs had to be adapted for three main reasons, which are elaborated on in Chapter 2:

1) Prediction year — The CMS-HCC risk adjustment model uses a prospective, rather than a concurrent model;

2) Population — The CMS-HCCs were developed using data from the aged (age ≥ 65) or disabled (age < 65) Medicare populations, as compared to the private individual, small group primarily under age 65 population; and

3) Type of spending — The CMS-HCCs are configured to predict medical spending excluding outpatient prescription drug spending as compared to medical and prescription drug

[8] In 2012, the Version 12 CMS-HCC risk adjustment model was used for Medicare Part C capitation payments for most managed care plans and the Version 21 model was used for PACE plans. Because the Version 21 model was recently developed following an extensive clinical revision and reclassification process, it was used as the starting point for creating the HHS-HCC classification and is used in this paper as the comparison CMS-HCC model. The CMS-HCC model has since been revised—the Version 22 CMS-HCC model will be used for Part C risk adjustment beginning in 2014.

spending. We call the revised clinical classification that is the basis of HHS risk adjustment the HHS-HCC clinical classification.

Separate Adult, Child, and Infant Models. In addition to revising the Medicare CMS-HCC clinical classification to be applicable to the largely under-age-65 individual and small group markets, we considered subpopulation differences *within* the HHS risk adjustment population. Clinical reasoning and empirical investigation led us to conclude that separate adult (age 21+), child (age 2-20), and infant (age 0-1) models are desirable for the risk adjustment population.

Plan Liability versus Total Expenditures. To account for differences in plan actuarial risk across actuarial value levels, we considered plan liability and total expenditure risk scores. A person's total expenditure risk score is the same regardless of plan, because expected total expenditures do not vary based on a plan's cost sharing rules. In contrast, a plan liability risk score predicts the medical expenditures that a plan is actually liable for, given its actuarial value and cost sharing structure. It incorporates the predicted effect of both health status and plan cost sharing on expected plan liability. An individual has a different plan liability risk score depending on the metal tier of the plan.

The plan liability risk score cannot be obtained by simply multiplying a person's total expenditure risk score by his/her plan's actuarial value because the amount plans pay is not constant as expenditures increase (i.e., it is non-linear, primarily because of the presence of deductibles). We instead develop model plan benefits at each actuarial value level (including deductible, coinsurance, and out-of-pocket limits), array each enrollee's health expenditures against those model benefits to derive estimated plan expenditures, and use these estimated expenditures to develop separate plan liability models on the same population for each actuarial value level. We use each enrollee's plan selection to determine which model to use when creating the enrollee's plan liability risk score.

Induced Demand Due to Cost-sharing Reductions. We also considered how to address the potential higher utilization among individuals who are enrolled in cost sharing reduction plans. A direct adjustment in the risk adjustment model for induced demand due to cost sharing reductions was not possible due to lack of the required data in the risk adjustment model calibration sample. As an alternative, a multiplicative adjustment to the risk score was developed. We chose to account for induced demand associated with more generous actuarial value of cost sharing reduction plans in the risk adjustment model because premiums for cost sharing reduction plans are required to be the same for all actuarial value levels of cost sharing reduction plans (in contrast to differing metal levels, where premiums can vary). For the Medicare Advantage program, induced demand due to lower cost sharing for Medicare-Medicaid dual eligible beneficiaries are adjusted for directly in the risk adjustment model by including a risk factor for dual eligible status.[9] Similarly, for the Medicare Part D program, induced demand due to lower cost sharing for low-income beneficiaries are adjusted for directly in the risk adjustment model by calibrating separate models for low-income beneficiaries.

[9] Medicare Advantage has recently proposed separate risk adjustment models by Medicare-Medicaid dual eligibility status. For details see:
https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2017.pdf

1.3.2 Risk Adjustment Payment Transfer Formula

The transfer formula uses the output of the risk adjustment model – plan liability risk scores – as an input to the transfer formula. The goal of risk adjustment transfers is to account for health risk differences while preserving permissible premium differences. Transfers are not intended to reflect costs due to differences in scope of coverage or costs that can be reflected in permissible rating differences. The payment transfer formula averages all individual risk scores in risk adjustment covered plans, and uses the plan liability risk scores, combined with other factors,[10] to calculate the funds transferred between plans. The payment transfer formula is based on the difference between two plan premium estimates: 1) premium with risk selection, and 2) premium without risk selection. Transfers are intended to bridge the gap between these two premium estimates, that is, to account for health risk difference while preserving permissible premium differences. If the difference between the two premium estimates is positive, a plan receives a risk adjustment payment. If the difference is negative, a plan is "charged" and owes a risk adjustment charge. The payment transfer formula is discussed in greater detail in Chapter 4.

[10] These other factors include plan allowable premium rating, actuarial value, induced demand, geographic costs, market share, and the Statewide average premium.

CHAPTER 2: HHS-HCC DIAGNOSTIC CLASSIFICATION

A diagnostic classification system provides the diagnostic framework for developing a risk adjustment model that uses patient diagnoses and demographic information to predict medical spending. This chapter describes the HHS-HCC diagnostic classification, how it was developed, and how the HHS-HCCs were selected and grouped for the HHS risk adjustment model. Because the CMS-HCC diagnostic classification (Pope et al., 2004) was used as a starting point to develop the HHS-HCC diagnostic classification, we provide an overview of that system as well.

2.1 Principles of Risk Adjustment

Determining which diagnosis codes should be included, how they should be grouped, and how the diagnostic groupings should interact for risk adjustment purposes was a critical step in the development of the HHS risk adjustment model. The following 10 principles, discussed in the HHS Notice of Benefit and Payment Parameters for 2014 proposed rule (proposed 2014 Payment Notice) (77 Federal Register 73118), guided the creation of the CMS-HCC diagnostic classification system as well as the HHS-HCC diagnostic classification system:

> ***Principle 1*—Diagnostic categories should be clinically meaningful**. Each diagnostic category is a set of International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) or International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes.[11] These codes should all relate to a reasonably well-specified disease or medical condition that defines the category. Conditions must be sufficiently clinically specific to minimize opportunities for gaming or discretionary coding. Clinical meaningfulness improves the validity and interpretability of the classification system.
>
> ***Principle 2*—Diagnostic categories should predict medical (including drug) expenditures**. Diagnoses in the same HCC should be reasonably homogeneous with respect to their effect on both current (this year's) costs (concurrent risk adjustment) or future (next year's) cost (prospective risk adjustment).
>
> ***Principle 3*—Diagnostic categories that will affect payments should have adequate sample sizes to permit accurate and stable estimates of expenditures**. Diagnostic categories used in establishing payments should have adequate sample sizes in available data sets. The data cannot reliably determine the expected cost of extremely rare diagnostic categories.
>
> ***Principle 4*—In creating an individual's clinical profile, hierarchies should be used to characterize the person's illness level within each disease process, while the effects of unrelated disease processes accumulate**. Because each new medical problem adds to an individual's total disease burden, unrelated disease processes should increase

[11] ICD-10-CM replaced ICD-9-CM effective October 1, 2015. CMS sought comment on a draft crosswalk in August 2015 and posted a final ICD-10-CM diagnostic classification crosswalk on the CCIIO website on October 19, 2015 at https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/DIY-tables-1092015.xlsx.

predicted costs of care. However, the most severe manifestation of a given disease process principally defines its impact on costs. Therefore, related conditions should be treated hierarchically, with more severe manifestations of a condition dominating (and zeroing out the effect of) less serious ones.

***Principle 5*—The diagnostic classification should encourage specific coding**. Vague diagnostic codes should be grouped with less severe and lower-paying diagnostic categories to provide incentives for more specific diagnostic coding.

***Principle 6*—The diagnostic classification should not reward coding proliferation**. The classification should not measure greater disease burden simply because more diagnosis codes are present. Hence, neither the number of times that a particular code appears, nor the presence of additional, closely related codes that indicate the same condition should increase predicted costs.

***Principle 7*—Providers should not be penalized for recording additional diagnoses (monotonicity)**. This principle has two consequences for modeling: (1) no HCC should carry a negative payment weight, and (2) a condition that is higher-ranked in a disease hierarchy (causing lower-rank diagnoses to be ignored) should have at least as large a payment weight as lower-ranked conditions in the same hierarchy. (There may be exceptions, as when a coded condition represents a radical change of treatment of a diseases process.)

***Principle 8*—The classification system should be internally consistent (transitive)**. If diagnostic category A is higher-ranked than category B in a disease hierarchy, and category B is higher-ranked than category C, then category A should be higher-ranked than category C. Transitivity improves the internal consistency of the classification system and ensures that the assignment of diagnostic categories is independent of the order in which hierarchical exclusion rules are applied.

***Principle 9*—The diagnostic classification should assign all ICD-9-CM or ICD-10-CM codes (exhaustive classification)**. Because each diagnostic code potentially contains relevant clinical information, the classification should categorize all ICD-9-CM or ICD-10-CM codes.

***Principle 10*—Discretionary diagnostic categories should be excluded from payment models**. Diagnoses that are particularly subject to intentional or unintentional discretionary coding variation or inappropriate coding, or that are not clinically or empirically credible as cost predictors, should not increase cost predictions. Excluding these diagnoses reduces the sensitivity of the model to coding variation, coding proliferation, gaming, and upcoding.

In designing the diagnostic classification, principles 7 (monotonicity), 8 (transitivity), and 9 (exhaustive classification) were generally followed. For example, if the expenditure weights for the models did not originally satisfy monotonicity, constraints were imposed to create models that did. Judgment was used to make tradeoffs among other principles. For example, clinical meaningfulness (principle 1) is often best served by creating a very large number of detailed clinical groupings. But a large number of groupings conflicts with adequate sample sizes for

each category (principle 3). Another tradeoff is encouraging specific coding (principle 5) versus predictive power (principle 2). In current coding practice, nonspecific codes are common. If these codes are excluded from the classification system, predictive power may be sacrificed. We approached the inherent tradeoffs involved in designing a classification system using empirical evidence on frequencies and predictive power; clinical judgment on relatedness, specificity, and severity of diagnoses; and professional judgment on incentives and likely provider responses to the classification system.

2.2 Diagnostic Reclassification: CMS-HCC → HHS-HCC

2.2.1 Medicare CMS-HCC Diagnostic Classification[12]

When we developed the HHS-HCC diagnostic classification system in 2012, we began with the CMS-HCC diagnostic classification system.[13] The HHS-HCC risk adjustment model uses health plan enrollee diagnoses (and demographics) to predict medical expenditure risk. To obtain a clinically meaningful and statistically stable system, the tens of thousands of ICD-9-CM codes (and beginning October 1, 2015, ICD-10-CM codes) used to capture diagnoses must be grouped into a smaller number of organized categories that produce a diagnostic profile of each person. The diagnostic classification is key in determining the ability of a risk adjustment model to distinguish high from low cost individuals. The classification also determines the sensitivity of the model to intentional or unintentional variations in diagnostic coding, an important consideration in real-world risk adjustment.

The CMS-HCC diagnostic classification system begins by classifying all ICD-9-CM diagnosis codes into Diagnostic Groups, or DXGs (see ***Figure 2.1***). Each ICD-9-CM code maps to exactly one DXG, which represents a well-specified medical condition or set of conditions, such as the DXG for *Type II Diabetes with Ketoacidosis or Coma*. DXGs are further aggregated into Condition Categories (CCs). CCs describe a broader set of similar diseases. Although they are not as homogeneous as DXGs, diseases within a CC are related clinically and with respect to cost. An example is the CC for *Diabetes with Acute Complications*, which includes in addition to the DXG for *Type II Diabetes with Ketoacidosis or Coma*, also the DXGs for *Type I Diabetes* and *Secondary Diabetes* (each with ketoacidosis or coma).

Hierarchies are imposed among related CCs, so that a person is coded for only the most severe manifestation among related diseases. After imposing hierarchies, CCs become Hierarchical Condition Categories (HCCs). For example, diabetes diagnosis codes are organized in the Diabetes hierarchy, consisting of three CCs arranged in descending order of clinical severity and cost, from 1) *Diabetes with Acute Complications* to 2) *Diabetes with Chronic Complications* to 3) *Diabetes without Complication*. Thus, a person with diagnosis code of

---

[12] For a primer on the CMS-HCC risk adjustment model, see section 2 in Pope, Kautter, Ingber et al. (2011) at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/downloads/evaluation_risk_adj_model_2011.pdf

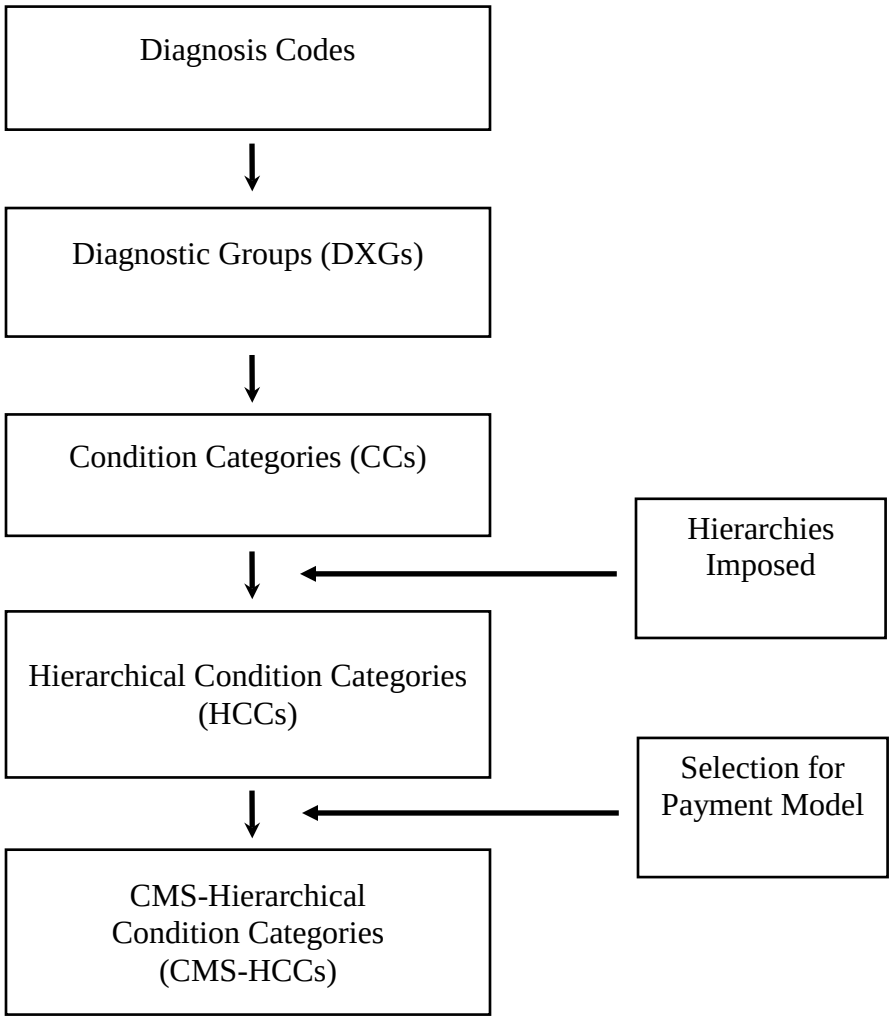
[13] In 2012, the Version 12 CMS-HCC risk adjustment model was used for Medicare Part C capitation payments for most managed care plans and the Version 21 model was used for PACE plans. Because the Version 21 model was recently developed following an extensive clinical revision and reclassification process, it was used as the starting point for creating the HHS-HCC classification and is used in this paper as the comparison CMS-HCC model. The CMS-HCC model has since been revised—the Version 22 CMS-HCC model will be used for Part C risk adjustment beginning in 2014.

Diabetes with Acute Complications is excluded from being coded with *Diabetes with Chronic Complications* and is also excluded from being coded with *Diabetes without Complication*. Similarly, a person with a diagnosis code of *Diabetes with Chronic Complications* is excluded from being coded with *Diabetes without Complication*. Although HCCs reflect hierarchies among related disease categories, for unrelated diseases, HCCs accumulate, i.e., the model is "additive." For example, a female with both *Rheumatoid Arthritis* and *Breast Cancer* has (at least) two separate HCCs coded, and her predicted cost will reflect increments for both conditions.

Because a single individual may be coded for none, one, or more than one HCC, the CMS-HCC model can individually price tens of thousands of distinct clinical profiles. The model's structure thus provides, and predicts from, a detailed comprehensive clinical profile for each individual.

The CMS-HCC Version 21 prospective risk adjustment model included the 87 HCCs (out of a total of 201 HCCs) that best predict future (next year's) Medicare Part A and Part B expenditures.

Figure 2.1
Hierarchical Condition Categories Aggregations of Diagnosis Codes



Three characteristics of the CMS-HCC classification system required review and adaptation for use with the HHS risk adjustment model:

1. Population — The CMS-HCCs were developed using data from the aged or disabled Medicare population.[14] Although every diagnosis code was mapped and categorized into a diagnostic grouping, for some conditions (such as pregnancy and neonatal complications) the sample size in the Medicare population is quite low. With larger sample sizes in the commercial population, HCCs were re-examined to study the homogeneity of costs in groups as well as the level of costs for infant, child, and adult subpopulations under age 65.

2. Type of Spending — The CMS-HCCs are configured to predict medical spending. A separate classification, the Medicare Prescription Drug Hierarchical Condition

[14] By aged we mean beneficiaries currently eligible for Medicare by age, and by disabled we mean beneficiaries currently eligible for Medicare by disability.

Categories (RxHCC) classification, is used to predict Medicare Part D drug spending. The HHS-HCCs predict the sum of medical and drug spending.

3. Prediction Year — The CMS-HCC classification is primarily designed for use with a prospective risk adjustment model, using base year diagnoses and demographic information to predict the next year's spending. The HHS risk adjustment model is concurrent, using current year diagnoses and demographics to predict the current year's spending. Medical conditions may have different implications in terms of current year costs and future costs; HCC groupings should reflect those differences.

To begin the reclassification process, we examined empirical results of the CMS-HCC full classification applied to 2009 and 2010 commercial population data.

2.2.2 Systematic Review with Clinical Consultants

During 2011 and 2012, CMS and our contractor worked with clinicians over a period of several months to systematically review and adapt the CMS-HCC classification for use with the primarily under-age-65 commercial population. The principles of risk adjustment (section 2.1) were presented to clinicians and discussed at the beginning of the review process.

The clinical review consisted of structured discussions in which the HCCs and their component diagnostic groups and diagnosis codes were discussed in sets by hierarchy (body system or disease group). The clinicians reviewed the MarketScan® empirical data (see section 3.3) for the full sample and three subsamples—adults, children, and infants—including sample size, mean expenditures, total spending regression coefficients (estimated incremental contribution to total spending), and predictive ratios (accuracy of prediction for diagnostic groups). The consultants provided clinical input in terms of:

- Clinical interpretation of empirical results;
- Clinical similarities and differences of specific diseases;
- Diagnosis and treatment differences by subpopulation (e.g., infants, children, adolescents, pregnant women);
- Severity and chronicity of illness, including cost implications in a concurrent versus prospective model;
- Criteria, discretion, and variability in diagnosis;
- Potential changes in diagnosis coding due to upcoming implementation of ICD-10-CM diagnosis codes, and to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5);[15] and
- Whether specific diagnoses or disease groups were overly discretionary and subject to "gaming."

[15] Although the final version of DSM-5 was not released until May 2013, draft versions were posted online in 2012 during the reclassification process.

Clinicians' input informed the HHS-HCC diagnostic classification and the selection of payment model HHS-HCCs.

2.2.3 HHS-HCC Diagnostic Reclassification

Integrating our empirical analyses, clinical input, and further background research, we split or reconfigured several HCCs, expanding the number of HCCs in the full classification from 201 CMS-HCCs to 264 HHS-HCCs. Many HCCs were split to better predict costs within disease groups, such as those in the metabolic, blood, psychiatric, and injury hierarchies.

2.3 Selection and Grouping of Payment HHS-HCCs

There are 264 HHS-HCCs in the full diagnostic classification, but only a subset is used for risk adjustment purposes. They are termed payment HHS-HCCs. Determining which HHS-HCCs would be included in the risk adjustment model was an iterative process, involving input from contractors, clinicians, specialist advisors, and CMS staff.

2.3.1 Criteria for Selection

In addition to the established criteria for selecting payment HCCs used in the Medicare models (based on the principles of risk adjustment), we had further parameters designed to support stability in the markets, especially in the initial years of market reform implementation. Following a series of discussions, we used the following criteria for selecting payment HHS-HCCs. Payment model HHS-HCCs should:

- Represent clinically significant medical conditions with significant costs for the target population;
- Contain sufficient sample size for stable results;
- Exclude (or limit the impact of) diagnoses particularly subject to discretionary coding;
- Exclude diagnoses that represent poor quality of care;
- Identify chronic or systematic conditions that represent insurance risk selection or risk segmentation, rather than random acute events that represent insurance risk; and
- Apply only to the appropriate model age group (infant, child, adult).

We conducted a comprehensive review process, selecting 127 payment HHS-HCCs to be used in the risk adjustment model (**Table 2.1**). Consistent with the risk adjustment principles described previously, the HHS risk adjustment model excludes HHS-HCCs containing diagnoses that are vague or nonspecific (e.g., cough), discretionary in medical treatment or coding (e.g., attention deficit disorder), or not medically significant (e.g., heartburn). The payment model also excludes HHS-HCCs that do not add empirically to costs (e.g., non-melanoma forms of skin cancer).

Although there are clear criteria for inclusion or exclusion, that determination was especially challenging for some HHS-HCCs. As an example, substance use disorders and mental health conditions are chronic and, in some cases, costly conditions. Ensuring access to treatment for these conditions is an important public health objective and a key purpose of risk adjustment is to reduce incentives for plans to attract or avoid patients based on risk selection. Limiting facility and physician or counseling networks for behavioral health is an observed method that plans use to discourage enrollment (Barry, et al., 2012). At the same time, the possibility of "gaming" diagnoses through strategic coding behavior is a valid concern for these HHS-HCCs—these are diagnoses whose reported prevalence has responded strongly to coding incentives in the Medicare Advantage program. CMS ultimately chose to include the more severe forms of these conditions as HHS-HCCs. CMS issued a draft ICD-10 to HHS-HCC crosswalk on August 13, 2015 to solicit feedback from the public on the underlying codes prior to implementation. While commenters generally supported the crosswalk, we received requests to include less severe forms of the conditions included, which were also previously excluded as ICD-9 codes. Consistent with the principles discussed above, we included the more severe forms of conditions to maintain consistency with the original HHS-HCC development.

Table 2.1:
Risk Adjustment Payment Model HHS-HCCs

HCC Number	HCC Label
HHS_HCC001	HIV/AIDS
HHS_HCC002	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock
HHS_HCC003	Central Nervous System Infections, Except Viral Meningitis
HHS_HCC004	Viral or Unspecified Meningitis
HHS_HCC006	Opportunistic Infections
HHS_HCC008	Metastatic Cancer
HHS_HCC009	Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia
HHS_HCC010	Non-Hodgkin`s Lymphomas and Other Cancers and Tumors
HHS_HCC011	Colorectal, Breast (Age < 50), Kidney, and Other Cancers
HHS_HCC012	Breast (Age 50+) and Prostate Cancer, Benign/Uncertain Brain Tumors, and Other Cancers and Tumors
HHS_HCC013	Thyroid Cancer, Melanoma, Neurofibromatosis, and Other Cancers and Tumors
HHS_HCC018	Pancreas Transplant Status/Complications
HHS_HCC019	Diabetes with Acute Complications
HHS_HCC020	Diabetes with Chronic Complications
HHS_HCC021	Diabetes without Complication
HHS_HCC023	Protein-Calorie Malnutrition
HHS_HCC026	Mucopolysaccharidosis
HHS_HCC027	Lipidoses and Glycogenosis
HHS_HCC028	Congenital Metabolic Disorders, Not Elsewhere Classified
HHS_HCC029	Amyloidosis, Porphyria, and Other Metabolic Disorders
HHS_HCC030	Adrenal, Pituitary, and Other Significant Endocrine Disorders

HCC Number	HCC Label
HHS_HCC034	Liver Transplant Status/Complications
HHS_HCC035	End-Stage Liver Disease
HHS_HCC036	Cirrhosis of Liver
HHS_HCC037	Chronic Hepatitis
HHS_HCC038	Acute Liver Failure/Disease, Including Neonatal Hepatitis
HHS_HCC041	Intestine Transplant Status/Complications
HHS_HCC042	Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis
HHS_HCC045	Intestinal Obstruction
HHS_HCC046	Chronic Pancreatitis
HHS_HCC047	Acute Pancreatitis/Other Pancreatic Disorders and Intestinal Malabsorption
HHS_HCC048	Inflammatory Bowel Disease
HHS_HCC054	Necrotizing Fasciitis
HHS_HCC055	Bone/Joint/Muscle Infections/Necrosis
HHS_HCC056	Rheumatoid Arthritis and Specified Autoimmune Disorders
HHS_HCC057	Systemic Lupus Erythematosus and Other Autoimmune Disorders
HHS_HCC061	Osteogenesis Imperfecta and Other Osteodystrophies
HHS_HCC062	Congenital/Developmental Skeletal and Connective Tissue Disorders
HHS_HCC063	Cleft Lip/Cleft Palate
HHS_HCC064	Major Congenital Anomalies of Diaphragm, Abdominal Wall, and Esophagus, Age < 2
HHS_HCC066	Hemophilia
HHS_HCC067	Myelodysplastic Syndromes and Myelofibrosis
HHS_HCC068	Aplastic Anemia
HHS_HCC069	Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn
HHS_HCC070	Sickle Cell Anemia (Hb-SS)
HHS_HCC071	Thalassemia Major
HHS_HCC073	Combined and Other Severe Immunodeficiencies
HHS_HCC074	Disorders of the Immune Mechanism
HHS_HCC075	Coagulation Defects and Other Specified Hematological Disorders
HHS_HCC081	Drug Psychosis
HHS_HCC082	Drug Dependence
HHS_HCC087	Schizophrenia
HHS_HCC088	Major Depressive and Bipolar Disorders
HHS_HCC089	Reactive and Unspecified Psychosis, Delusional Disorders
HHS_HCC090	Personality Disorders
HHS_HCC094	Anorexia/Bulimia Nervosa
HHS_HCC096	Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes
HHS_HCC097	Down Syndrome, Fragile X, Other Chromosomal Anomalies, and Congenital Malformation Syndromes
HHS_HCC102	Autistic Disorder
HHS_HCC103	Pervasive Developmental Disorders, Except Autistic Disorder

HCC Number	HCC Label
HHS_HCC106	Traumatic Complete Lesion Cervical Spinal Cord
HHS_HCC107	Quadriplegia
HHS_HCC108	Traumatic Complete Lesion Dorsal Spinal Cord
HHS_HCC109	Paraplegia
HHS_HCC110	Spinal Cord Disorders/Injuries
HHS_HCC111	Amyotrophic Lateral Sclerosis and Other Anterior Horn Cell Disease
HHS_HCC112	Quadriplegic Cerebral Palsy
HHS_HCC113	Cerebral Palsy, Except Quadriplegic
HHS_HCC114	Spina Bifida and Other Brain/Spinal/Nervous System Congenital Anomalies
HHS_HCC115	Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy
HHS_HCC117	Muscular Dystrophy
HHS_HCC118	Multiple Sclerosis
HHS_HCC119	Parkinson`s, Huntington`s, and Spinocerebellar Disease, and Other Neurodegenerative Disorders
HHS_HCC120	Seizure Disorders and Convulsions
HHS_HCC121	Hydrocephalus
HHS_HCC122	Non-Traumatic Coma, and Brain Compression/Anoxic Damage
HHS_HCC125	Respirator Dependence/Tracheostomy Status
HHS_HCC126	Respiratory Arrest
HHS_HCC127	Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes
HHS_HCC128	Heart Assistive Device/Artificial Heart
HHS_HCC129	Heart Transplant
HHS_HCC130	Congestive Heart Failure
HHS_HCC131	Acute Myocardial Infarction
HHS_HCC132	Unstable Angina and Other Acute Ischemic Heart Disease
HHS_HCC135	Heart Infection/Inflammation, Except Rheumatic
HHS_HCC137	Hypoplastic Left Heart Syndrome and Other Severe Congenital Heart Disorders
HHS_HCC138	Major Congenital Heart/Circulatory Disorders
HHS_HCC139	Atrial and Ventricular Septal Defects, Patent Ductus Arteriosus, and Other Congenital Heart/Circulatory Disorders
HHS_HCC142	Specified Heart Arrhythmias
HHS_HCC145	Intracranial Hemorrhage
HHS_HCC146	Ischemic or Unspecified Stroke
HHS_HCC149	Cerebral Aneurysm and Arteriovenous Malformation
HHS_HCC150	Hemiplegia/Hemiparesis
HHS_HCC151	Monoplegia, Other Paralytic Syndromes
HHS_HCC153	Atherosclerosis of the Extremities with Ulceration or Gangrene
HHS_HCC154	Vascular Disease with Complications
HHS_HCC156	Pulmonary Embolism and Deep Vein Thrombosis
HHS_HCC158	Lung Transplant Status/Complications

HCC Number	HCC Label
HHS_HCC159	Cystic Fibrosis
HHS_HCC160	Chronic Obstructive Pulmonary Disease, Including Bronchiectasis
HHS_HCC161	Asthma
HHS_HCC162	Fibrosis of Lung and Other Lung Disorders
HHS_HCC163	Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections
HHS_HCC183	Kidney Transplant Status
HHS_HCC184	End Stage Renal Disease
HHS_HCC187	Chronic Kidney Disease, Stage 5
HHS_HCC188	Chronic Kidney Disease, Severe (Stage 4)
HHS_HCC203	Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism
HHS_HCC204	Miscarriage with Complications
HHS_HCC205	Miscarriage with No or Minor Complications
HHS_HCC207	Completed Pregnancy With Major Complications
HHS_HCC208	Completed Pregnancy With Complications
HHS_HCC209	Completed Pregnancy with No or Minor Complications
HHS_HCC217	Chronic Ulcer of Skin, Except Pressure
HHS_HCC226	Hip Fractures and Pathological Vertebral or Humerus Fractures
HHS_HCC227	Pathological Fractures, Except of Vertebrae, Hip, or Humerus
HHS_HCC242	Extremely Immature Newborns, Birthweight < 500 Grams
HHS_HCC243	Extremely Immature Newborns, Including Birthweight 500-749 Grams
HHS_HCC244	Extremely Immature Newborns, Including Birthweight 750-999 Grams
HHS_HCC245	Premature Newborns, Including Birthweight 1000-1499 Grams
HHS_HCC246	Premature Newborns, Including Birthweight 1500-1999 Grams
HHS_HCC247	Premature Newborns, Including Birthweight 2000-2499 Grams
HHS_HCC248	Other Premature, Low Birthweight, Malnourished, or Multiple Birth Newborns
HHS_HCC249	Term or Post-Term Singleton Newborn, Normal or High Birthweight
HHS_HCC251	Stem Cell, Including Bone Marrow, Transplant Status/Complications
HHS_HCC253	Artificial Openings for Feeding or Elimination
HHS_HCC254	Amputation Status, Lower Limb/Amputation Complications

2.3.2 Aggregate Grouping

To balance the competing goals of improving predictive power and limiting discretionary coding, as well as to create a risk adjustment model in which less experienced insurance carriers are not disadvantaged, a subset of payment HHS-HCCs were re-grouped into larger aggregate clusters. We re-grouped payment HHS-HCCs for the following reasons:

1. To reduce model complexity by limiting the number of effective payment HHS-HCCs;

2. To avoid HHS-HCCs with low sample size and possibly unstable high-cost estimates;

3. To limit upcoding by severity within an HCC hierarchy; and

4. To reduce additivity within disease groups (but not across disease groups) in order to decrease the sensitivity of the model to coding proliferation.

After the re-grouping was complete, the number of payment HHS-HCCs in the risk adjustment model was effectively reduced from 127 to 100.

2.4 Comparison of CMS-HCC Classification and HHS-HCC Classification

Table 2.2 provides summary statistics comparing the Version 21 CMS-HCC classification to the HHS-HCC classification, when it was initially developed, allowing for a side-by-side comparison of the payment HCCs for each risk adjustment model.

Table 2.2:
Summary statistics for CMS-HCC vs. HHS-HCC classifications

Model characteristic	Medicare V21 CMS-HCC model	Health exchange HHS-HCC model
Model type	Prospective	Concurrent
Number of HCCs in full classification	201	264
Number of payment HCCs	87	127
Number of payment HCCs after grouping[1]	—	100
Number of ICD-9-CM codes in full classification	14,445	14,445
Number of ICD-9-CM codes in payment model	3,124	3,439
Percent of ICD-9-CM codes in payment model	22%	24%

NOTES:

1. As previously discussed, in the HHS risk adjustment model, sets of clinically-related HCCs are grouped together as a single HCC with a single coefficient.

2. CMS, Centers for Medicare & Medicaid Services; HCC, Hierarchical Condition Categories; HHS, Health and Human Services; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; V, Version.

CHAPTER 3: HHS-HCC RISK ADJUSTMENT MODEL – DATA AND METHODS

3.1 Data Source

Because we did not have a data set for the individual and small group markets, we used a commercial dataset for the original calibration for the 2014 benefit year. The calibration sample for the HHS-risk adjustment model consisted of 2010 Truven MarketScan® Commercial Claims and Encounter data. The MarketScan® data is a large, widely used, nationally dispersed, proprietary database sourced largely from large employers and health plans. Employees, spouses, and dependents covered by employer-sponsored private health insurance are included. The MarketScan® sample includes enrollees from all 50 States and the District of Columbia. Although MarketScan® represents primarily the large employer commercial market rather than the small group/individual market, we believe it is the dataset that best approximates the likely relationship between diagnoses and relative expenditures for the individual and small group markets, holding constant the generosity of plan benefits (essential health benefits and metal level). We compared the age, sex, and regional distribution of the MarketScan® sample to the expected HHS risk adjustment population (Trish, Damico, and Claxton et al., 2011; Buettgens, Garrett, and Holahan, 2010). We found that overall they were similar, although the MarketScan® data had more children and fewer young adults, and more sample members in the South and fewer in the Northeast and West than the expected risk adjustment population.

3.2 Sample

An enrollee is included in the concurrent modeling sample if the enrollee has at least 1 month of enrollment, is enrolled in a preferred provider organization (PPO) or other fee-for-service (FFS) health plan[16], has no payments made on a capitated basis, has prescription drug coverage, and has integrated mental health/substance abuse coverage.[17] The primary goals of the sample selection criteria were to ensure that 1) enrollees had complete expenditure and diagnosis data, 2) enrollees included those entering (e.g., newborns) and exiting (e.g., decedents) enrollment during the year, and 3) enrollees had health care coverage similar to the essential health benefits under the Affordable Care Act.

We have continued to use these inclusion criteria for each year of MarketScan® data we have incorporated in the HHS risk adjustment models.

3.3 Expenditures

The HHS-HCC risk adjustment model predicts health care expenditures for which plans are liable, which exclude enrollee cost sharing. This is termed a plan liability risk adjustment model, which has been used in other payment systems such as Medicare Part C and Part D (Pope et al., 2004; Kautter et al., 2012). We considered predicting total expenditures and then adjusting to plan liability with a multiplicative plan actuarial value factor. However, this approach might not accurately capture plan liability levels due to the non-linear relationship of plan liability to total expenditures. Although there can exist more complex plan cost sharing designs, for illustrative purposes, we can approximate plan liability as follows. Plan

[16] Other fee-for-service health plans include, for example, indemnity, consumer directed, and high-deductible health plans.

[17] Additionally, mothers with bundled newborn claims, and newborns with no birth records, were excluded.

liability is zero percent of total expenditures below the applicable deductible, one minus the coinsurance percentage of total expenditures between the deductible and the out-of-pocket limit, and one hundred percent of total expenditures above the out-of-pocket limit.[18] Deductibles, and to a lesser extent coinsurance and out-of-pocket maximums, are anticipated to vary systematically with plan metal levels, being highest in catastrophic plans and lowest in platinum plans.

Using the MarketScan® inpatient, outpatient, and drug services files, we summed total payments (submitted charges – non-covered charges – pricing discounts), which include enrollee cost sharing. We then trended the MarketScan® expenditures to the applicable benefit year by applying a constant annual growth rate.[19] Once expenditures were trended, standardized benefit design parameters (deductibles, coinsurance rates, out-of-pocket limits) were applied for each metal level to simulate plan liability expenditures for each metal level. Plan liability expenditures were then annualized by dividing them by the fraction of months in the dataset year that each individual is enrolled in the plan (i.e., by the eligibility fraction). Annualized expenditures are the "per member per month" amount multiplied by 12. Note that annualized expenditures were not truncated.[20]

Finally, plan liability expenditures were converted to relative plan liability expenditures, which are defined as plan liability expenditures divided by a denominator. A relative plan liability expenditure of 1.0 corresponds to the average plan liability expenditure for the calibration sample. The denominator was calculated as follows. For the entire calibration sample, we calculated the mean plan liability for each metal level, and then took a weighted average of these means, where the weights were based on a forecasted distribution of enrollment in the projected benefit year across the five metal levels. Going forward, we use the term "plan liability" to mean "relative plan liability."

In short, we simulated plan liability expenditures for each metal level from total expenditures for each sample member (that is, we applied different benefit structures to the same sample). An alternative approach would have been to model actual plan liability (payments) for enrollees in MarketScan® plans grouped into metal tiers by the plans' actual actuarial values. However, MarketScan® provides sufficient plan benefit information to calculate plan actuarial value for only a small fraction of its sample. Also, grouping plans by actuarial value would have led to different samples of individuals for each metal level model estimation, which would have reduced sample sizes for each model and led to differences in unmeasured factors across metal level samples. Simulating plan liability on the full sample for each metal also means that (as intended) the model estimates for the HCCs do not compensate for differential induced demand across metals.

---

[18] See section 3.12 "Model Updates" for our incorporation of preventive services into our simulation of plan liability, beginning in 2017 benefit year risk adjustment.

[19] See section 3.12 "Model Updates" for a discussion of our change in growth rates for 2017 benefit year risk adjustment.

[20] See section 4.3 "High Risk Enrollee Pooling in HHS Risk Adjustment" for a discussion of how we could truncate expenditures in future recalibrations.

3.4 Demographics

The HHS risk adjustment model uses MarketScan® enrollee demographics and diagnoses to predict plan liability expenditures for each enrollee. The demographic factors employed are age and sex. Age is measured as of the last month of enrollment, which in general results in infants with age 0 having been born in the dataset year.[21] Age ranges were determined by the age distribution of the commercial population, as well as consideration of post 2014 market reform rules for the individual and small group markets. There are 18 age/sex categories for adults, and 8 age/sex categories for children. As described below, age/sex categories for infants are not used.[22] Adults are defined as ages 21+, children are ages 2-20, and infants are ages 0-1. The age categories for adult male and female are ages 21-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, and 60+. The age categories for children, male and female, are ages 2-4, 5-9, 10-14, and 15-20.

3.5 Diagnoses

Only diagnosis codes from sources allowable for risk adjustment when HHS is operating risk adjustment on behalf of a State are included in the diagnosis-level file. The goal of the restrictions on source of diagnoses is to improve the quality, accuracy, and auditability of diagnoses used for risk adjustment. For example, clinical laboratory diagnoses, which include "rule outs" and diagnoses not verified by a clinician, were excluded. Allowable diagnoses include those from inpatient hospital claims, outpatient facility claims (hospital outpatient, rural health clinic, Federally qualified health center, and community mental health clinic), and professional claims (diagnoses are generally not available on prescription drug claims, including for the MarketScan® data). In addition, diagnoses from outpatient facility claims and professional claims are restricted to those with at least one CPT/HCPCS code corresponding generally to face-to-face encounters with a clinician.[23]

3.6 Subpopulations

Due to the inherent clinical and cost differences in the adult, child, and infant populations, we developed separate risk adjustment models for each group. The adult and child models have similar specifications, with age/sex demographic categories and HCCs (individual HCCs and aggregate HCC groupings) predicting annualized plan liability expenditures.

However, infants have low frequencies for most HCCs leading to unstable parameter estimates in an additive model. Because of this, the infant model utilizes a categorical approach in which infants are assigned a birth maturity (by length of gestation and birth weight) or age 1 category, and a disease severity category (based on HCCs other than birth maturity). There are four Age 0 birth maturity categories--Extremely Immature; Immature;

---

[21] If an infant is measured as age 0 in the dataset year, but was born in the prior year, the infant's age is recoded to age 1. See section 3.12 "Model Updates" for details.

[22] There are age 0 male and age 1 male additive terms in the infant model.

[23] Additional information about allowable CPT/HCPCS codes and diagnoses in operations can be found in the "HHS-Developed Risk Adjustment Model Algorithm Instructions " and "Technical Details" tables for 2015 risk adjustment at https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/DIY-instructions-10-16-15.pdf and https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/DIY-tables-1092015.xlsx.

Premature/Multiples; Term--and an Age 1 maturity category. Age 0 infants are assigned to one of the four birth maturity categories and age 1 infants are assigned to the Age 1 maturity category. Sample counts and plan liability expenditures monotonically decrease when moving from higher to lower maturity categories.

There are 5 disease severity categories based on the clinical severity and associated costs of the non-maturity HCCs: Severity Level 5 (Highest Severity) to Severity Level 1 (Lowest Severity).[24] Examples of severity level assignments are:

- Level 5 -- HCC 137 (Hypoplastic Left Heart Syndrome and Other Severe Congenital Heart Disorders);
- Level 4 -- HCC 127 (Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes);
- Level 3 -- HCC 45 (Intestinal Obstruction);
- Level 2 -- HCC 69 (Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn); and,
- Level 1 -- HCC 37 (Chronic Hepatitis).

As expected, plan liability expenditures monotonically decrease when moving from higher to lower disease severity categories, and counts generally increase. All infants (age 0 or 1) are assigned to a disease severity category based on the single highest severity level of any of their non-maturity HCCs. HCCs not appropriately diagnosed for infants—such as pregnancy and psychiatric HCCs—were excluded from the infant disease severity categories. Infants with no severity HCCs are assigned to Level 1.

When cross-classified, the 5 maturity categories and 5 severity categories define 25 mutually-exclusive categories. Each infant is assigned to one of the 25 categories. Finally, there are two additive terms for sex, for age 0 males and age 1 males.[25]

3.7 Model Estimation

All risk adjustment models are estimated by weighted least squares regression.[26] The dependent variable is annualized simulated plan liability expenditures, and the weight is the person-specific sample eligibility fraction. Annualization and weighting—which is equivalent on an annual basis to predicting per member per month expenditures weighting by the number of

[24] In assigning HCCs to infant severity levels, the HCC hierarchies are maintained. If two HCCs are in a hierarchical relationship, the higher-ranking HCC is assigned to the same or a higher severity level than the lower-ranking HCC.

[25] Male infants have higher costs than female infants due to increased morbidity and neonatal mortality; the 25 mutually-exclusive categories presume female infants.

[26] We investigated various non-linear approaches to model estimation that might have been better able to account for the non-linearities in plan liability. However, these models suffer from several important shortcomings, including complexity, lack of transparency, and not predicting mean expenditures accurately for all diagnostic and demographic subgroups, or even for the overall sample. We concluded that, evaluated against a broad range of criteria for real-world risk adjustment, weighted least squares is the preferable estimation method.

months each individual is eligible for the sample--appropriately adjusts for months of enrollee eligibility in the sample. For the adult model, independent variables include 18 age/sex demographic categories, 114 HCC diagnosis groups, and 16 disease interactions (discussed below), and for the child model, 8 age/sex demographic categories and 119 HCC diagnosis groups. For the infant model independent variables include 25 categories defined by birth maturity for age 0, age 1, and diagnostic severity, and 2 age/sex demographic additive terms.

In each adult and child regression model, we include a binary indicator variable for each individual HCC that is not included in an aggregate HCC grouping. In addition, we include a binary indicator for each aggregate HCC grouping. In the latter case it indicates whether or not the enrollee had at least one HCC in the aggregate HCC grouping.

In addition, we impose equality constraints on certain of the individual HCC incremental predicted costs if the unconstrained incremental predicted costs would violate the principle that higher-clinically-ranked conditions in an HCC hierarchy should have higher predicted costs. Constraints generally have the effect of averaging two or more groups together when, unconstrained, there is a violation of clinical logic.[27]

3.8 Disease Interactions

For the adult models, including disease interaction terms better reflected plan liability across metal levels and improved model performance.[28] Based on empirical findings, as well as clinical review, we developed a set of eight diagnostic markers of severe illness: HCC 2 (Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock); HCC 42 (Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis); HCC 120 (Seizure Disorders and Convulsions); HCC 122 (Non-Traumatic Coma, Brain (Compression/Anoxic Damage)); HCC 125 (Respirator Dependence/Tracheostomy Status); HCC 126 (Respiratory Arrest); HCC 127 (Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes); and, HCC 156 (Pulmonary Embolism and Deep Vein Thrombosis). A severe illness indicator variable was defined as having at least one of the eight diagnostic markers of severe illness.[29]

The severe illness indicator was interacted with individual HCCs and aggregate HCC groupings.[30] The disease interactions that met minimum sample size and incremental predicted expenditure thresholds were included in the model. The incremental predicted expenditures for the disease interactions were categorized into medium and high cost categories. For each category, we included a binary indicator variable in the regression model for whether or not the enrollee had at least one disease interaction in the category. Finally, a hierarchy was imposed such that if an enrollee was in the high cost disease interaction category, the enrollee was excluded from the medium cost category. In sum, a person can have at most one disease interaction coefficient or incremental predicted expenditure. This constraint was imposed

---

[27] In addition, to increase the stability of the transplant HCC coefficients in the child risk adjustment model, we impose equality constraints for a selected group of these HCCs in the child model. See section 3.12 "Model Updates" for details.

[28] Disease interactions were empirically unimportant in the child model and were not included. The infant model is a categorical model.

[29] The diagnostic markers of severe illness are also included in the model not interacted with other diagnoses (HCCs).

[30] High frequency, high incremental expenditure disease interactions tended to include severe illnesses.

because clinical reasoning and empirical evidence indicated that a single one of the diagnostic markers sufficed to distinguish the most severely ill patients among those with the underlying interacted diagnosis.

3.9 Predicted Plan Liability Expenditures

For an enrollee in a given metal level plan, the total predicted plan liability expenditures is the sum of the incremental predicted plan liability expenditures (coefficients) from the relevant metal level model. For adults and children, this is the sum of the age/sex, HCC, and disease interaction coefficients.[31] For infants, this is the sum of the maturity/disease-severity category and additive sex coefficients, if male.

As discussed earlier, plan liability expenditures were converted to relative plan liability expenditures, resulting in a relative plan liability expenditure of 1.0 for the average plan liability expenditure in the calibration sample. Converting "actual" plan liability expenditures to relatives automatically converts "predicted" plan liability expenditures to relatives. Thus a "predicted" relative plan liability expenditure of 1.0 corresponds to the average "predicted" plan liability expenditure for the calibration sample. Going forward, we use the term "predicted plan liability" to mean "predicted relative plan liability."

3.10 Risk Score Calculation

Below we provide an example of how the risk adjustment model output is applied to calculate individual risk scores and the "plan liability risk score (PLRS)" that is used in the calculation of transfer payments and charges.[32] In the HHS methodology, the risk score for an enrollee is defined as the total predicted relative plan liability expenditures for the enrollee based on the relevant HHS-HCC risk adjustment model for the enrollee's age group and plan metal level. For the relevant metal level of the enrollee's plan, the total predicted relative plan liability expenditures are calculated as follows. For an adult (age 21+), it is the sum of the age/sex, HCC, and disease interaction risk factors; for a child, it is the sum of the age/sex and HCC risk factors,[33] and for infants, it is the sum of the appropriate maturity/disease severity category and male additive term.

An adjustment to the risk score is made to the risk score for enrollees in individual market cost-sharing plan variations in Marketplaces because individuals who qualify for cost sharing reductions may utilize health care services at a higher rate than would be the case in the absence of cost sharing reductions. This adjustment for induced demand due to cost-sharing reductions is multiplicative, and is applied to the risk score. Because premiums for all cost-sharing reduction plan variations are required to be the same despite the increased actuarial value of coverage, we account for the induced demand associated with cost-sharing plan variations as part of the risk adjustment model, and not as part of the risk transfer formula.

---

[31] The child risk adjustment models do not have disease interactions.

[32] When the plan average PLRS is calculated, *all* plan enrollees are counted in the numerator of the risk transfer formula, but only *billable* plan enrollees (parents and the three oldest children) are counted in the denominator. For details, see Chapter 5.

[33] Recall that disease interactions were not empirically important for the child HHS-HCC risk adjustment models.

The induced demand utilization factor was determined based on an analysis of the expected difference in expenditures for enrollees in Qualified Health Plans of different actuarial values. For this analysis, the Actuarial Value Calculator was used. The induced utilization factors appear to be broadly consistent with results from the RAND Health Insurance Experiment (Newhouse, 1984). The cost-sharing reduction adjustment factors used in the HHS risk adjustment methodology are below in **Table 3.1**.

Table 3.1:
Cost-Sharing Reduction Adjustment

Household Income	**Plan AV**	**Induced Utilization Factor**
Silver Plan Variant Recipients		
100-150% of FPL	Plan Variation 94%	1.12
150-200% of FPL	Plan Variation 87%	1.12
200-250% of FPL	Plan Variation 73%	1.00
>250% of FPL	Standard Plan 70%	1.00
Zero Cost-Sharing Recipients		
<300% of FPL	Platinum (90%)	1.00
<300% of FPL	Gold (80%)	1.07
<300% of FPL	Silver (70%)	1.12
<300% of FPL	Bronze (60%)	1.15
Limited Cost-Sharing Recipients		
>300% of FPL	Platinum (90%)	1.00
>300% of FPL	Gold (80%)	1.07
>300% of FPL	Silver (70%)	1.12
>300% of FPL	Bronze (60%)	1.15

Table 3.2 provides illustrative examples of the PLRS calculation, using the final 2017 risk adjustment coefficients and assuming a silver metal level plan. Enrollee 1 is male and aged 56, without cost-sharing reductions, with two chronic conditions – diabetes with complications, and congestive heart failure. Predicted relative plan liability expenditures for these demographic and diagnostic risk factors in the adult silver model are 0.429, 0.925, and 3.095, respectively. Therefore his total predicted relative plan liability expenditures is 4.449, and since he is not entitled to cost-sharing reductions (that is, the induced utilization factor is 1.00), his PLRS is 4.449. Enrollee 2 is female and aged 11, with asthma. Her total predicted relative plan liability expenditures from the child silver model is 0.316 (0.085+0.231). However, she also is a zero cost-sharing plan variation enrollee, so that her total predicted expenditures is multiplied by her induced utilization factor 1.12, resulting in a PLRS of 0.354. Enrollee 3 is male and aged 0, with a term birth and severity level 1. His total predicted plan liability expenditures from the infant silver model is 1.380 (0.772+0.608), and since he doesn't have cost sharing reductions, it is his PLRS as well.

Table 3.2:
Plan Liability Risk Scores for Silver Metal Level Plan -- Illustrative Examples
(2017 Risk Adjustment)

Enrollee	Predicted relative plan liability expenditures	Induced demand factor	Plan liability risk score
Enrollee 1			
Age 56 and male	0.429	—	—
Diabetes with complications	0.925	—	—
Congestive heart failure	3.095	—	—
Total	4.449	1.00	4.449
Enrollee 2			
Age 11 and female	0.085	—	—
Asthma	0.231	—	—
Total	0.316	1.12	0.354
Enrollee 3			
Age 0 and male	0.608	—	—
Term and severity level 1	0.772	—	—
Total	1.380	1.00	1.380

NOTE: Plan liability risk score equals the total predicted relative plan liability expenditures based on the relevant HHS-HCC risk adjustment model for the enrollee's age group and plan's metal level, multiplied by the induced demand utilization factor due to cost sharing reductions.

3.11 Model Evaluation

The predictive accuracy of a risk adjustment model for individuals is typically judged by the percentage of variation in individual expenditures explained by the model (as measured by the R-squared statistic). To test the performance of the HHS-HCC risk adjustment models for subgroups, we calculate the expenditure ratio of predicted to actual weighted mean plan liability expenditures, which is commonly termed the "predictive ratio." If prediction is perfect, mean predicted expenditures will equal mean actual expenditures, and the predictive ratio will be 1.00. As a rule of thumb, predictive ratios with a margin of error of 10 percent in either direction ($0.90 \leq$ predictive ratio ≤ 1.10) indicate reasonably accurate prediction.

3.12 Model Updates Since 2014 Risk Adjustment

Since finalizing the 2014 HHS risk adjustment methodology in the 2014 Payment Notice, we have made several updates to the risk adjustment methodology through notice and comment rulemaking to improve the accuracy and consistency of the program.

3.12.1 – 2015 Benefit Year

For 2015 risk adjustment, our primary goal was to maintain stability, given the new market and program. In the 2015 Payment Notice final rule (79 Federal Register 13744), we discussed how we would incorporate premium assistance Medicaid alternative plans in the HHS risk adjustment methodology for 2014 risk adjustment. We finalized that we would use the same factor that we use to adjust for induced utilization for individuals enrolled in cost-sharing plan variations to adjust for induced utilization for individuals enrolled in the corresponding Medicaid alternative plan variations. We also sought comment on how to best adjust the geographic cost factors or geographic rating areas in future years to address potential premium distortions in less populous rating areas.[34] Based on comments received, commenters did not support making additional adjustments to the geographic cost factor, stating that the time and resources needed to calculate and implement such an adjustment would be considerable, and that any such adjustment would be unlikely to have a material impact on final risk adjustment results. We did not adjust the geographic cost factors or geographic rating areas, but noted that we would monitor 2014 risk adjustment data for any potential premium distortions.

3.12.2 – 2016 Benefit Year

In the 2016 Payment Notice final rule (80 Federal Register 10750), we recalibrated the risk adjustment models for the first time, to provide risk adjustment factors that better reflect more recent treatment patterns and costs. We considered using a single year of more recent MarketScan® data, similar to the use of only 2010 MarketScan® data for 2014 risk adjustment. However, given our desire for stability in the initial years and concerns about small sample sizes for some conditions, we ultimately finalized blending, or averaging, coefficients using separately solved models of 2011, 2012, and 2013 MarketScan® data for 2016 benefit year risk adjustment. Additionally, we finalized a change in how we categorize age 0 infants who do not have birth codes. We stated in operational guidance, also known as the "Do It Yourself (DIY)" software, that infants without birth codes would be assigned an ''Age 0, Term'' factor in risk adjustment operations.[35] We did so under the assumption that issuers paid the birth costs, yet the birth HCCs were missing (perhaps because claims were bundled with the mother's, whose claims were excluded). Upon further analysis of age 0 and age 1 claims, we found that age 0 infants without birth HCCs had costs more similar to age 1 infants by severity level. We finalized that these infants should be assigned to age 1 by severity level. For many age 0 infants without birth HCCs, the birth could have occurred in the prior year or was paid for by a different issuer. We finalized making this change in 2016 benefit year operations. In the 2016 recalibration samples, we also finalized constraining six transplant status HCC coefficients (other than kidney) in the child model, as the sample sizes of transplants are smaller in the child than the adult model. Because the levels and changes in the child transplant relative coefficients appeared to be dominated by random instability, we believe the accuracy of the model is improved by constraining these coefficients. We continue to monitor the child transplant relative coefficients, and will adjust them if needed in future recalibrations.

---

[34] For details on the transfer formula, including geographic cost and rating factors, see Chapter 5.

[35] HHS-Developed Risk Adjustment Model Algorithm "Do It Yourself (DIY)" Software Instructions https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/DIY-instructions-10-16-15.pdf

3.12.3 – 2017 Benefit Year

Finally, in the 2017 Payment Notice final rule (81 Federal Register 12204), we finalized a similar approach to recalibration as we did for the 2016 benefit year. We updated the underlying data using the three most recent years of data (2012, 2013, and 2014 MarketScan®) and blended the coefficients from three separately solved datasets to derive the 2017 risk adjustment model factors (see Appendix A). We also finalized incorporating preventive services into our simulation of plan liability in the recalibration of the risk adjustment models for the 2017 benefit year. We identified preventive services for the 2012, 2013, and 2014 MarketScan® samples using procedure and diagnosis codes, prescription drug therapeutic classes, and enrollee age and sex. We relied on lists of preventive services from several major issuers, the preventive services used for the AV Calculator, and Medicare's preventive services benefit to operationalize preventive services definitions for incorporation in the risk adjustment models. We then adjusted plan liability by adding 100 percent of preventive services covered charges to simulate plan liability for all metal levels. We also applied standard benefit cost sharing rules by metal level to covered charges for non-preventive services. Total adjusted simulated plan liability is the sum of preventive services covered charges, and non-preventive services simulated plan liability.

We re-estimated the risk adjustment models by metal level, predicting plan liability adjusted to account for preventive services without cost sharing. We compared the model coefficients predicting original (that is, non-adjusted for preventive services) and adjusted simulated plan liability. Adjusting for preventive services increases age-sex coefficients relative to HCC coefficients, especially in the lower metal tiers (bronze and silver), and in age/sex ranges with higher preventive services expenditures (for example, young adult females). The implication of the changes to the model coefficients is that the risk scores of healthy enrollees (whose risk scores are based solely on model age-sex coefficients) will likely rise relative to the risk scores of the less healthy (whose risk scores include one or more HCC coefficients in addition to an age-sex coefficient), especially in bronze and silver plans. As a result of the risk score changes for individuals, we expect that the incorporation of preventive services will increase the risk scores of bronze and silver plans with healthier enrollees relative to other plans' risk scores when preventive services are taken into account. This incorporation of preventive services will more accurately compensate risk adjustment covered plans with enrollees who use preventive services.

We also sought comment on how to account for partial year enrollment in the risk adjustment methodology, incorporating prescription drugs as risk factors in the risk adjustment model, and addressing the data lag to better account for high cost conditions that may have new treatments. Based on commenters' feedback on the need to better model the risk of high-cost conditions and rapidly changing health care costs, we re-examined the underlying trend factor we used to trend medical and prescription drug expenditures in the MarketScan® data, because those expenditures account for a large portion of the recent changes in costs to treat high-cost conditions. Because we were using the same trend for both sets of expenditures, we looked at historical MarketScan® drug data, subdivided by traditional (including branded and generic) drugs, specialty drugs, and medical and surgical expenditures, and found varying growth rates. In order to address commenters' feedback, we consulted with actuaries and industry reports to derive a specialty drug trend rate and traditional drug trend rate through the 2017 benefit year. We believe that using these more granular trend rates better reflect the growth in specialty drug

expenditures and drugs generally as compared to medical and surgical expenditures. Further, we believe that more accurately trending drug expenditures through the 2017 benefit year will more accurately compensate issuers providing new treatments associated with specific HCCs by providing a more finely tuned estimate of the relative costs of various conditions under the HHS risk adjustment methodology. We finalized the incorporation of different trend factors for (i) traditional drugs, (ii) specialty drugs, and (iii) medical and surgical expenditures for 2017 risk adjustment in the 2017 Payment Notice final rule.

CHAPTER 4: RISK ADJUSTMENT MODEL IMPROVEMENTS

We have recently received feedback from the public suggesting, among other things, that our risk adjustment model does not capture the risk associated with partial year enrollment, that it undercompensates new or fast-growing plans, that it is based on outdated data, and that it would be improved by including prescription drug utilization data as a predictor. We sought comment on these areas for improvement in the 2017 Payment Notice (81 Federal Register 12204) and received many comments in support of exploring these areas further. In this chapter, in response to these comments and as a result of our continuing evaluation of potential data sources and the risk adjustment methodology, we explore six topics: 1) partial year enrollment, 2) prescription drug utilization as a predictor in the model, 3) pooling of high cost enrollees, 4) an evaluation of concurrent and prospective risk adjustment models, 5) data for 2018 recalibration, and 6) data for 2019 recalibration. We seek comment on these topics. Any changes to the HHS risk adjustment methodology will be implemented through future notice and comment rulemaking. In the longer term, we would like to explore the possibility of using socioeconomic status or other sociodemographic factors as predictors in the risk adjustment model. However, we believe there are other short-term improvements that can be made in advance of that undertaking. Our continuing priorities as we consider improvements to our methodology will be to ensure that the provisions we incorporate in recalibration complement each other in improving the accuracy and performance of the HHS risk adjustment model in a data-driven fashion, while balancing the need for model predictability and stability.

4.1 Partial Year Enrollment

After the 2014 benefit year of risk adjustment, we received feedback that some issuers experienced higher than expected claims costs for partial year enrollees for the initial year of the risks adjustment program. We also received feedback that some of the public believe the methodology does not capture enrollees with chronic conditions who may not have accumulated diagnoses in their partial year enrollment. On the other hand, compared to full year enrollees of the same relative risk, partial year enrollees are less likely to have spending that exceeds the deductible or annual limitation on cost sharing. Commenters have stated that enrollees with partial year enrollments of 6 months or less yielded high medical loss ratios (MLRs) and financial losses for issuers.

We sought comment in the 2017 Notice of Benefit and Payment Parameters on how the risk adjustment methodology could be made more accurate for partial year enrollees. We received comments generally supporting addressing partial year enrollees in the risk adjustment model. One commenter noted that many medical events for enrollees in the commercial market (for example, maternity, surgeries) represent acute rather than chronic events, such that the enrollee may incur most of their annual medical expenses during a short period of time. Commenters also suggested that the use of prescription drug claims could help capture a partial year enrollee with a chronic condition who does not have a provider encounter with a documented diagnosis. We received feedback suggesting that the impact of partial year enrollment could be measured by taking a population that had multiple years of enrollment and comparing risk scores and health care costs when only a partial year is considered. Commenters noted that Massachusetts' alternate risk adjustment methodology for its State-operated risk adjustment program includes a duration adjustment for partial year enrollment, and suggested

that HHS consider additional analysis to determine whether a similar approach is appropriate for the HHS risk adjustment methodology. Commenters requested member-level adjustments or duration adjustments. Commenters stated that unverified special enrollment periods have produced selection issues for health plans, as enrollees enter through a special enrollment period, utilize high-cost services, and then switch to a lower metal level plan in the following open enrollment period or drop coverage altogether. Some commenters opposed an explicit adjustment for partial year enrollees, because they said such an adjustment would accommodate liberal enforcement of special enrollment periods, incentivizing issuers to employ loose eligibility standards to gain members, but ultimately eroding individual market stability. One commenter cautioned that any additions to the model to account for partial year enrollment should improve reliability and predictive power, not influence clinical judgment or plan behavior with respect to enrollees' coverage.

In general, we believe that individual and small group health plans whose enrollees are representative of the market will be risk adjusted accurately under the HHS risk adjustment methodology. However, risk adjustment may be inaccurate when a plan's enrollees differ substantially from the market as a whole with respect to characteristics that are not adjusted for in the risk adjustment model. For example, if a plan has an unrepresentative enrollee population by enrollment duration, and risk associated with enrollee duration is not fully captured through other aspects of our methodology, then for that plan, partial year enrollment is not accurately accounted for in the HHS risk adjustment methodology. Specifically, if the risk adjustment methodology does not fully capture risk for partial year enrollment, then if the plan had higher than average enrollment duration, the plan risk score might be too high, and similarly, if the plan had lower than average enrollment duration, the plan risk score might be too low. We noted commenters' requests to review the use of additional factors, such as in Massachusetts' alternative risk adjustment methodology, and the use of wholly separate models that account for duration of enrollment and metal level. Below we discuss our findings. However, we note that partial year enrollees and enrollment durations in the Marketscan® data are likely different from those in the individual and small group markets.

Initially, we looked at the predicted expenditures, actual expenditures, and predictive ratios (that is, the ratios of predicted to actual weighted mean plan liability expenditures) by enrollment duration (ED) group (1 month, 2 months, …, 12 months) annualized for 2014 MarketScan® adults in our risk adjustment concurrent modeling sample. If the prediction is perfect, mean predicted expenditures will equal mean actual expenditures, and the predictive ratio will be 1.00. A rule of thumb is that predictive ratios with a margin of error of 10 percent in either direction ($0.90 \leq$ predictive ratio ≤ 1.10) indicate reasonably accurate prediction. [36] Table 4.1 below displays the predictive ratios of the overall adult concurrent sample. The data in this table indicates that actuarial risk for all adult enrollees with short enrollment periods tends to be underpredicted, and for adult enrollees with full enrollment periods (12 months) tends to be slightly overpredicted (Table 4.1). One potential explanation for these results is that enrollees will tend to be coded with HCCs for expensive, acute events (e.g., Opportunistic Infections) when they have the high cost, acute event. For enrollees with full enrollment, the costs of these expensive, acute events are spread out over the entire risk adjustment year. However, for

[36] A margin of error of 10 percent is obviously arbitrary, but we are only suggesting a rule of thumb here.

enrollees with partial year enrollment, those costs will largely be concentrated in a shorter period.

Table 4.1:
Predicted and Actual Expenditures by Enrollment Duration
Adults in the Concurrent Sample

	Mean Months Enrollment	Ratio (Predicted/Actual)				
		Platinum	Gold	Silver	Bronze	Catastrophic
ED_01	1	0.598	0.592	0.559	0.515	0.514
ED_02	2	0.668	0.669	0.652	0.623	0.622
ED_03	3	0.742	0.748	0.745	0.726	0.726
ED_04	4	0.792	0.799	0.800	0.788	0.787
ED_05	5	0.828	0.834	0.839	0.834	0.833
ED_06	6	0.853	0.859	0.864	0.862	0.862
ED_07	7	0.869	0.872	0.875	0.874	0.874
ED_08	8	0.929	0.933	0.939	0.940	0.940
ED_09	9	0.945	0.946	0.948	0.947	0.947
ED_10	10	0.950	0.949	0.947	0.945	0.945
ED_11	11	0.952	0.949	0.943	0.937	0.937
ED_12	12	1.021	1.020	1.020	1.021	1.021

We then examined the 2014 MarketScan® adult silver model, but with enrollment duration (1 month, 2 months, …, 11 months—12 months is the reference group and therefore is not included) binary indicator variables added as additional risk factors. Below in Table 4.2, are the resulting estimated enrollment duration factors.

Table 4.2:
Adult Silver Model Enrollment Duration Factors

Label	Parameter Estimate
Enrollment duration 1 month	0.424
Enrollment duration 2 months	0.368
Enrollment duration 3 months	0.275
Enrollment duration 4 months	0.227
Enrollment duration 5 months	0.196
Enrollment duration 6 months	0.174
Enrollment duration 7 months	0.175
Enrollment duration 8 months	0.101
Enrollment duration 9 months	0.092
Enrollment duration 10 months	0.098
Enrollment duration 11 months	0.111
Enrollment duration 12 months (reference group)	0.000

When partial year enrollment is included in the model as a duration factor, the impact is relatively small. The factors themselves did not appear to reflect noticeably higher costs associated with partial year enrollees, perhaps because the factors incorporate the risk of both partial year enrollees with no payment HCCs and partial year enrollees with payment HCCs, two populations with risk effects that tend to offset each other.

We then calculated full, separate 2014 MarketScan® adult silver models based on enrollment periods (months 1-4, months 5-8, and months 9-12). Table 4.3 below lists selected parameter estimates for the 2014 MarketScan® adult silver risk adjustment separate models by enrollment duration. Specifically, we calculated the difference between the HCC coefficients in the months 1-4 model and the months 9-12 model, and identified the top 20 HCCs with the greatest differences in coefficients. These HCCs tend to be those that represent expensive, acute events during the risk adjustment year.

Other things being equal, we believe that separate models by enrollment duration are preferred, since they will predict accurately by enrollment duration subpopulations overall, by age/sex categories, and by HCC disease groups. While separate models for partial year enrollment indicate that certain diagnoses' predicted expenditures are directly affected by enrollment length, one concern is that separate models for all risk adjustment coefficients may present false precision in predicting the costs associated with some conditions, particularly conditions with small sample sizes. In addition, separate models would add to the complexity of the HHS risk adjustment methodology, which for each risk adjustment year is currently calibrating 45 risk adjustment models (3 MarketScan® data years x 3 age groups x 5 metal levels = 45 risk adjustment models) to develop the 15 blended coefficients risk adjustment models that are used (coefficients are blended across the 3 MarketScan® data years). If separate models by partial year enrollment were developed, this would increase the number of models that are required to be calibrated for a given risk adjustment year. As noted above, the use of enrollment duration adjustment factors in the model appeared to have relatively little impact on an enrollee's risk score. However, when we created separate risk adjustment models based on enrollment duration, we found in many cases very different coefficients for expensive, acute conditions by duration as compared to chronic conditions, which appeared to be relatively stable. We note that these model results may not fully reflect the experience of some commenters, because we continue to conduct our modeling on a commercial dataset, with largely employer plans, which may not reflect the unique enrollment duration and health status of the individual and small group markets.

Table 4.3:
2014 MarketScan® Adult Silver Risk Adjustment Models by Enrollment Duration: 1-4 Months, 5-8 Months, and 9-12 Months – Top 20 HCCs with largest changes between 1-4 Month and 9-12 Month Models

HCC Number	HCC Label	1-4 Month Model	5-8 Month Model	9-12 Month Model
HHS_HCC128_14	Heart Assistive Device/Artificial Heart	96.94	82.02	32.30
HHS_HCC129_14	Heart Transplant	96.94	82.02	32.30

HCC Number	HCC Label	1-4 Month Model	5-8 Month Model	9-12 Month Model
HHS_HCC125_14	Respirator Dependence/Tracheostomy Status	90.26	52.11	30.13
HHS_HCC251_14	Stem Cell, Including Bone Marrow, Transplant Status/Complications	70.61	43.34	26.54
HHS_HCC006_14	Opportunistic Infections	39.22	12.99	6.64
HHS_HCC034_14	Liver Transplant Status/Complications	45.52	17.76	16.06
HHS_HCC041_14	Intestine Transplant Status/Complications	57.70	89.33	30.06
HHS_HCC066_14	Hemophilia	68.31	67.90	41.03
HHS_HCC096_14	Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes	27.72	19.44	1.17
HHS_HCC121_14	Hydrocephalus	30.32	12.84	4.47
HHS_HCC067_14	Myelodysplastic Syndromes and Myelofibrosis	37.34	22.14	12.15
HHS_HCC068_14	Aplastic Anemia	37.34	22.14	12.15
HHS_HCC002_14	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	32.63	14.61	8.35
HHS_HCC042_14	Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis	33.72	14.35	10.43
HHS_HCC154_14	Vascular Disease with Complications	29.26	13.29	6.76
HHS_HCC145_14	Intracranial Hemorrhage	30.29	9.75	7.90
HHS_HCC003_14	Central Nervous System Infections, Except Viral Meningitis	27.63	12.34	5.99
HHS_HCC135_14	Heart Infection/Inflammation, Except Rheumatic	25.65	9.58	5.96
HHS_HCC004_14	Viral or Unspecified Meningitis	23.27	4.30	4.04
HHS_HCC126_14	Respiratory Arrest	24.69	11.38	8.58

One additional option we are considering is a hybrid approach combining enrollment duration adjustment factors and separate models. We are evaluating the feasibility of using enrollment duration factors in the model that would be interacted with individual HCCs and/or groupings of HCCs, where selected HCCs are determined by their sensitivity to the separate, partial year models' predicted parameters or coefficients. We are in the process of examining the parameters and results of this method, where the groupings are based on HCCs' sensitivity to enrollment duration (1-4 months, 5-8 months, 9-12 months).

We continue to evaluate the most appropriate way to account for partial year enrollees in the risk adjustment models, without inadvertently discouraging issuers from retaining enrollees or misrepresenting risk adjustment coefficients as a result of small sample sizes.

4.2 Prescription Drug Modeling

This section discusses adding prescription drug utilization as risk markers to the HHS-HCC model, to create a "hybrid" drug-diagnosis risk adjustment model. Section 4.2.1 discusses benefits of and concerns about adding drug utilization to a diagnosis-based risk adjustment model such as the HHS-HCC. Section 4.2.2 specifies criteria for evaluating risk adjustment models that incorporate prescription drugs. Section 4.2.3 develops an empirical framework for adding drug information to the HHS-HCC model. Section 4.2.4 describes the drug classification and aggregation systems used in this analysis. Section 4.2.5 identifies criteria for selecting drug-diagnosis pairs for the hybrid risk adjustment model and presents empirical and clinical analyses that helped select initial drug-diagnosis pairs for inclusion in the hybrid model. Section 4.2.6 presents estimates of illustrative hybrid risk adjustment models that use both diagnoses and drug utilization as risk markers. Section 4.2.7 provides recommendations, and identifies further anticipated development work on a hybrid risk adjustment model.

4.2.1 Benefits of and Concerns about Adding Prescription Drug Utilization to the HHS-HCC Risk Adjustment Model

In designing a risk adjustment model that incorporates prescription drug utilization, it is important to consider how such a model can improve risk adjustment that is based on diagnoses and age/sex, which are the risk markers used in the current HHS-HCC model. This section describes the major benefits to adding drugs to a diagnosis-based model to create a "hybrid" model, as well as some of the drawbacks or concerns.

Benefits

Imputing Missing Diagnoses. According to feedback we received after the initial year of risk adjustment, one of the most important roles for drug information is to fill in the gaps where diagnoses may be missing due to under-recording in medical claims or encounter data. In any one-year period, many patients with a chronic condition will have at least one claim that includes their diagnosis. However, for some patients, these diagnoses may be missing from a year's worth of claims. Clinicians may fail to enter the condition on a patient's chart during every office visit, for example, or there may be a stigma associated with certain health conditions that leads providers not to record it on claims. Clinical diagnostic reporting systems may focus on the diagnosis under current treatment, and not record the full set of a patient's underlying conditions. Also, not all individuals are enrolled in health plans for an entire year. They may lack physician visits during the time they are enrolled, but may fill prescriptions treating chronic conditions during even a short enrollment period. In these situations, drug utilization data may capture the existence of some conditions (diagnoses associated with drug treatment) that are missing in diagnoses entered on claims. As indicators of actual treatment provided, therefore, drug data can usefully augment incomplete diagnostic data from claims or encounters. Augmenting reported diagnoses with drug utilization may arguably lead to a more uniform and comprehensive capture of medical conditions across health plans with varying degrees of capture of conditions through diagnostic reporting.

Severity Indicator for a Specific Diagnosis. A second potential improvement that drug utilization patterns can offer is a more complete picture of the severity of illness. The HHS-HCCs already capture information about illness severity from diagnoses, but drugs can

potentially measure the severity of illness within a single HCC. A patient may receive first, second, or third lines of treatment that indicate increasing levels of severity. For example, the hybrid drug-diagnosis model presented in this paper includes Class IB and Class III antiarrhythmic medications, which are typically reserved for more serious cases due to these drugs' side effects. This type of differentiation could be a useful predictor of non-drug expenditures.

More Timely, Standardized Data. Drug data can be available more quickly than diagnoses from medical claims, is often more complete, and is often easier to access.[37] For example, there is often a lag between when an inpatient hospitalization ends and when the inpatient medical claim is submitted to and processed by the payer. Prescription drug claims may also be a more complete reflection of health conditions for patients whose chronic conditions do not require frequent physician visits. In addition, prescription drug data are standardized, and do not vary with provider coding patterns for diagnoses.

Mitigates the Financial Disincentive to Prescribe Expensive Medications. To the extent that a risk adjustment model incorporates prescription drug utilization, it will compensate plans that cover high-cost medications for its enrollees. This reduces the incentive for plans to restrict access to these medications, and is fairer to plans that enroll many people who require expensive drugs. Prescription drug risk adjustment models can also encourage plans to include more drugs in their formularies.

Concerns

Risk adjustment models based on diagnostic information have been widely used for payment in public programs such as Medicare and Medicaid for several years. As such, they are well understood and generally accepted among policymakers and the public. Models that use drug information for risk adjustment are less common in public programs, however, and their implementation raises several issues and concerns.[38]

Gaming, Perverse Incentives, and Discretionary Prescribing. A primary concern is the susceptibility of drug models to gaming and, more broadly, perverse incentives.[39,40] Gaming occurs if a drug is prescribed in order to trigger a higher payment from a risk adjustment model. Drug models may be particularly susceptible to this sort of gaming when there are inexpensive drugs included in therapeutic classes that are statistically linked to high total medical expenditures; in these situations, a small cost to the insurance plan (reimbursement for the drug)

---

[37] Hall, M.A. 2011. Risk Adjustment under the Affordable Care Act. A Guide for Federal and State Regulators. Issue Brief. Commonwealth Fund.

[38] Risk adjustment models that incorporate prescription drugs as additional risk factors are used in some public programs, for example, some State Medicaid agencies use the CDPS+Rx risk adjustment model (for details see http://cdps.ucsd.edu/).

[39] Winkelman, R. and Meymud, S. April 20, 2007. A Comparative Analysis of Claims-Based Tools for Health Risk Assessment. Society of Actuaries.

[40] Diagnosis-based models are not immune to gaming. For example, it has been well documented that Medicare Advantage health plans have engaged in diagnosis coding intensity to maximize their risk adjustment payments from CMS. See for example "Kronick and Welch, 2014. Measuring Coding Intensity in the Medicare Advantage Program. Medicare & Medicaid Research and Review, V4, N2, E1-E19."

can bring a relatively large increase in revenue. Although this same concern can be raised for diagnostic models in which the cost of coding an additional condition may be low, the prospect of medically unnecessary prescriptions raises added concern about wasteful or possibly even harmful treatments.

Perverse incentives arise in any risk adjustment model in which utilization indicators (such as prescriptions) trigger additional payments. In this case, treatment decisions may be influenced or distorted by financial considerations. Even for drugs with relatively specific clinical indications, there will typically be patients on the margin of the clinical indications, and financial incentives may influence prescribing for these marginal patients. Basing risk adjustment on drug utilization will tend to bias health plans towards drug rather than non-drug treatments, and towards use of the specific drug therapeutic classes linked to payment in risk adjustment. Plans will see a lower return (or even a negative return) from managing drug utilization, and may lessen their efforts to control drug costs.

Addressing the "gameability" of risk adjustment models that use drug information requires analysis to determine which drug classes (or individual drugs) are most susceptible, and how to devise groupings that strike a reasonable balance between predictive accuracy and reducing "gameability." This way, the model could be designed to minimize susceptibility to gaming and poor incentives. However, it is not clear that very many drug classes are non-discretionary. Substantial uncertainty or disagreement across providers over the circumstances in which drugs should be prescribed appears to exist for many drugs. Even anti-retroviral therapy for HIV/AIDS is now sometimes used prophylactically and may not indicate the presence of an HIV diagnosis. Hybrid risk adjustment models that add drugs to diagnoses are likely to gain most of their incremental accuracy in predicting health care expenditures from simply identifying individuals taking expensive drugs, even if this utilization is discretionary. It is not clear that drug utilization is less discretionary than other types of health utilization predictive of expenditures, such as hospitalizations for chronic conditions.

Sensitivity of Risk Adjustment to Variations in Prescription Drug Utilization. Aside from intentional gaming in response to risk adjustment incentives, incorporating drug utilization makes risk adjustment sensitive to variations in drug utilization patterns that exist for reasons other than enrollee health status. Many factors affect drug utilization, and with the inclusion of prescription drugs, risk adjustment would reflect these factors. Health plans with lower prescribing rates would incorrectly appear to have healthier populations, and would pay higher risk charges or receive lower risk payments.

For example, cost sharing and drug utilization management features differ across plans. Other things equal, drug utilization is expected to be lower in plans with higher cost sharing (such as bronze or silver plans) and aggressive drug utilization management, such as prior authorization, step therapy, quantity limits, restrictive formularies, and more stringent requirements to qualify for coverage of expensive (e.g., specialty) drugs. In general, plans providing greater access to prescription drug treatments, which will tend to include higher-actuarial-value plans such as gold and platinum, and higher premium plans within metal tier, will appear to have sicker populations but may not reflect the true health status of their enrollees.

Providers in different health plan provider networks may have different proclivities for using drug versus non-drug treatments for a medical condition (diagnosis), or use a different mix

of drug therapeutic classes to treat a condition. There may be different norms of physician prescribing behavior across areas or specialties, and the availability of physician specialties with different prescribing patterns may differ across areas. Individuals enrolled in different health plans may vary in their propensity to utilize drugs. For example, individuals planning to take expensive medications may be more likely to enroll in higher actuarial value plans such as gold or platinum plans, or other plans that provide cheaper access to drugs. Lower-income enrollees may use drugs at a lower rate, holding health status constant, because cost sharing for drugs or the associated physician visits are financial barriers.

Added Administrative Burden, Complexity, and Costs. Adding drug data to a diagnosis-based model introduces operational complexities. Additional data requirements increase the administrative burden associated with calibrating and applying the model. Clinical indications for drugs can change quickly, which would require frequent updates to the model specification and possibly the therapeutic classification groupings as well. As the model is calibrated before the start of the benefit year, it may be difficult to assess all updates or upcoming utilization pattern changes. Issuers of plans subject to risk adjustment would be required to report prescription drug utilization as well as diagnoses, and audit and verification of the reported data would be necessary.

Availability of Outpatient Drug Data Only. Another potential problem with drug models is that available data includes only outpatient prescription information. For a severity model, the omission of drugs provided in a hospital setting may introduce bias; in particular, hospitalized patients may appear to be less severely ill, because their drug utilization is not included in the model. However, often these patients will receive refills of their prescriptions after discharge, especially if the drugs are treating a chronic condition. If an imputation model were used, a diagnosis would be more likely to be on an inpatient claim, rendering the outdated drug data superfluous.

Multiple Indications for Most Drugs. Few drug classes are indicated for only one medical condition. Many drug classes are widely prescribed "off label" for indications that are not FDA-approved. Utilization of such drug classes can have very different implications for health care expenditures depending on the reasons for which they are prescribed. Presence of a drug class may not discriminate between high and low cost individuals if it is used for both high and low cost conditions. Some drug classes may be used for both diagnoses that have been included in the HHS HCC model, and diagnoses that have been intentionally excluded, making it problematic to maintain this distinction in a hybrid drug-diagnosis risk adjustment model. Specific drugs within a drug class may have varying indications; the utilization of such drug classes may not unambiguously indicate the presence of a specific diagnosis. The lack of clear, one-to-one associations between most drug classes and diagnoses makes development of a hybrid risk adjustment model that incorporates and integrates drug and diagnosis risk markers challenging.

4.2.2 Criteria for Evaluating Risk Adjustment Models Incorporating Prescription Drug Utilization

The criteria we are using to evaluate "hybrid" models including both diagnoses and prescription drugs follow from the benefits and concerns expressed in section 4.2.1. This section briefly discusses these criteria.

Criterion 1: Clinical/Face Validity

The hybrid model should have clinical face validity in the relationship it expresses between the risk markers (diagnoses, drugs) and health care expenditures, and in the relationship between drugs and associated diagnoses. The model should incorporate clinically salient and significant drug-diagnosis pairs where there is a clear relationship between a drug class and a single clinical diagnosis.

Criterion 2: Empirical/Predictive Accuracy

Drugs added to the model should increase the model's accuracy in predicting total health care expenditures, including drug and non-drug expenditures. Predictive accuracy should be improved overall and for specific subpopulations, for example for people with different severity levels of a disorder. Adding drugs to the model should improve the completeness of measurement of health plan enrollee health status by identifying the medical conditions or health states of individuals that are not captured by diagnoses. A more complete measurement of enrollee health status should improve the fairness of risk adjustment across all health plans.

Criterion 3: Incentives for Prescription Drug Utilization

If we determine that drugs should be added to the model, we would do so in a way that minimizes incentives for inappropriate use of prescription drugs, or over-prescription of drugs to maximize risk transfers. Models that reduce the incentive to over-prescribe drugs are preferred. At the same time, models should not discourage needed and appropriate use of drugs based on enrollees' clinical indications.

Criterion 4: Sensitivity to Variations in Prescription Drug Utilization

The hybrid model should not be overly sensitive to variations in discretionary prescription drug utilization. The model should incorporate variations in drug utilization that measure differences in health plan enrollee health status, not variation due to other factors, such as variation in physician prescribing patterns unrelated to patient health conditions.

Criterion 5: Incentives for Diagnosis Reporting

The effect of a hybrid model on the incentives for diagnosis reporting should be considered. Accurate and complete diagnosis reporting should not be discouraged by, for example, reducing predicted expenditures/risk scores when additional diagnoses are appropriately reported.

4.2.3 A Framework for Analyzing Added Prescription Drug Utilization to the HHS-HCC Risk Adjustment Model

In this section, we develop a framework for analyzing whether to add prescription drug utilization to the HHS-HCC model. Our goal is not to develop a final or comprehensive hybrid diagnosis/drug model, but to review issues and considerations, illustrate some possible approaches, examples, and preliminary findings, and solicit public feedback on the material presented.

The two main conceptual approaches we have identified to using drug information in a "hybrid" risk adjustment model are:

- Using drug data to impute (substitute for) missing diagnoses; and
- Using drug data as a severity indicator for a specific diagnosis.

With both of these approaches – imputation and severity – each drug class that enters the model specification would be chosen based on its relationship to a specific diagnosis or group of diagnoses. Linking drug utilization to specific diagnoses can be challenging: many drug therapeutic classes are used for multiple clinical conditions, the diagnostic indications for drugs can change and expand or contract over time, and physicians can prescribe drugs "off label" as they deem medically appropriate.

Statistical Predictive Power Approach. Given the complexity of linking drugs to diagnoses, another way to augment risk adjustment models with drug data is to include drug classes in the model on purely statistical grounds. This approach has the advantage of requiring little clinical consideration beyond the creation of clinically- and cost-coherent therapeutic classes. Some drugs may be markers for frailty or poor health in general, and a diagnosis-agnostic approach would allow the risk adjustment model to accommodate these linkages. Many of these drugs treat general symptoms such as pain, nausea, or anxiety that may be associated with many underlying diagnoses. The main disadvantage of this approach is that by omitting clinical consideration from the model specification, it makes interpretation of model coefficients difficult and the model may have less clinical face validity. We have emphasized the approaches linking drug classes to diagnoses in our development work, but we have utilized information from the pure statistical approach to help identify candidate drug classes to add to the HHS-HCC model. We welcome feedback on both approaches.

Implications of Conceptual Approaches for Empirical Modeling Specification

In the current diagnosis-only HHS-HCC model, an incremental predicted cost is associated with each additional diagnostic category (HCC) added to a person's diagnostic profile, subject to grouping of diagnoses into categories of related diagnoses (HCCs), diagnosis hierarchies (more severe HCCs exclude less severe related HCCs), and HCC constraints (some groups of HCCs are constrained to have a single incremental effect). Each conceptual approach to adding drugs to risk adjustment – creating a "hybrid" diagnosis/drug model – implies a different empirical model design and specification.

Under a pure imputation hybrid model approach, the predicted incremental total cost (drug and non-drug) for individuals with a given health condition is the same regardless of how the health condition is identified (whether by diagnosis or by drug utilization). The model would therefore be specified such that the predicted cost increment is equal for someone with a drug indicator only, or "drug only," as for someone with a diagnosis indicator only, or "diagnosis only" – and also equal for someone with both indicators.

Under a severity approach, the utilization of a drug class by itself would not be associated with any incremental predicted costs. Only if the drug class and a specific diagnosis are both present will the model predict incremental costs beyond the diagnosis alone. So a drug indicator only would predict zero incremental costs, a diagnosis indicator only would predict positive

incremental costs, and both a diagnosis and its related drug indicating severity within diagnosis would predict the highest incremental costs.

A third approach, the Rx dominant approach, incorporates elements of both the imputation and severity approaches. The rationale for this approach is that when a drug is utilized, the associated diagnosis is imputed, so the predicted expenditures should be the same irrespective of whether the diagnosis is reported. That is, the presence of the diagnosis does not add any information to the information from the drug being utilized. This model predicts the same incremental cost for the (i) drug only and (ii) both drug and diagnosis groups, but a different incremental cost (typically lower) for the diagnosis only group. The Rx dominant approach effectively establishes a hierarchy of clinical severity in which the drug indicator hierarchically excludes the diagnosis indicator. That is, individuals taking the drug are assumed to be more severely ill—whether or not they have the diagnosis marker—than individuals not taking the drug who have only the diagnosis marker.

Flexible/Generalized Empirical Framework. All of our conceptual approaches – severity, imputation, and Rx dominant – can be implemented as special cases of a generalized, flexible empirical model design. This generalized framework is itself a fourth conceptual approach that we consider. In the general formulation of this hybrid drug-diagnosis framework, each drug-diagnosis pair enters the model with three indicator variables:

- A diagnosis indicator (HCC), whose coefficient predicts the incremental costs associated with individuals who have the diagnosis;
- A drug class indicator (RXC), whose coefficient predicts the incremental costs associated with individuals who received a prescription in that drug class;
- An interaction indicator (RXC*HCC), whose coefficient predicts the incremental costs associated with individuals who have both the diagnosis and the drug prescription.

This general, flexible framework allows different predictions for the diagnosis only, drug only, and both diagnosis and drug groups. Variations or constraints on this framework can limit the model to estimate via imputation only, a severity relationship, Rx dominant, or other theoretical relationships between the drug class and its associated health condition(s). For example, in a pure imputation model, the predicted incremental expenditures are equal regardless of whether the individual has a diagnosis, a drug flag, or both. A severity model, on the other hand, predicts one level of expenditures for someone with the diagnosis only, a higher level for someone with the diagnosis and the associated drug class, and no additional expenditures for someone with the drug only. See Appendix B for a more complete description of the generalized econometric model and how constraints on the generalized model produce empirical specifications for the specific approaches. We welcome feedback from the public on these frameworks or any other frameworks we should consider to analyze incorporating drugs into the HHS-HCC model.

4.2.4 Drug Classification System: Prescription Drug Categories (RXCs)

To develop hybrid drug-diagnosis risk adjustment models, we need a clinically and empirically cohesive set of drug classes of a manageable number to augment the model specification. This chapter describes our Prescription Drug Categories (RXC) classification system, how it was developed, and how the RXCs were selected and grouped for our RXC risk adjustment model.

Each prescription drug is assigned a National Drug Code (NDC) maintained by the U.S. Food and Drug Administration (FDA). There are over 190,000 NDCs, which include prescription drugs as well as over-the-counter medications. NDC codes are contained in prescription drug claims data. Due to the large number of individual NDCs, it is necessary to use a therapeutic classification system that classifies individual NDCs into aggregated categories of related drugs used for similar therapeutic purposes, or having similar pharmacological properties.

The RXCs are based on the American Hospital Formulary Service (AHFS) Pharmacologic-Therapeutic Classification©, which is published by authority of the Board of the American Society of Health-System Pharmacists®. We chose to use the AHFS classification because it is widely used, widely available,[41] transparent, comprehensive, and regularly updated. While the United States Pharmacopeia (USP) classification is used for assessing compliance with the Essential Health Benefits (EHB), we did not believe it was best suited for this purpose because of the greater level of detail AHFS offers (in most therapeutic categories), and instances of difficulty with mapping NDCs to USP drug classes. The more frequent updates associated with AHFS is also making it easier to keep the prescription drug models up to date. NDC codes are classified into 527 AHFS therapeutic classes (TCs) at various levels of detail, from the broadest (2 digits) to the greatest detail (8 digits). We further aggregated these AHFS-TCs into 127 RXCs, 61 of which are designated as "payment" RXCs that are potentially suitable for inclusion in a payment risk adjustment model.

The following 10 principles guided the creation of the RXC classification system:

***Principle 1*—RXC categories should be clinically meaningful**. Each RXC is composed of a set of NDCs. These codes should all relate to a reasonably well-specified pharmacologic, therapeutic or chemical characteristic that defines the category. RXCs must be sufficiently clinically specific to minimize opportunities for gaming or discretionary coding. Clinical meaningfulness improves the face validity of the classification system to clinicians and its interpretability.

***Principle 2*—RXCs should predict total medical and drug expenditures**. NDCs in the same RXC should be reasonably homogeneous with respect to their effect on current year costs.

***Principle 3*—RXCs that will affect payments should have adequate sample sizes to permit accurate and stable estimates of expenditures**. RXCs used in establishing payments should have adequate sample sizes in available data sets. The data cannot reliably determine the expected cost of extremely rare categories.

[41] Although not in the public domain.

***Principle 4*—In creating an individual's clinical profile, hierarchies should be used to characterize the person's illness level within each RXC where appropriate, while the effects of unrelated prescriptions accumulate**. Because each new medical problem adds to an individual's total disease burden, unrelated prescriptions should increase predicted costs of care. However, the most severe manifestation of a given disease process principally defines its impact on costs. Therefore, related RXCs should be treated hierarchically, with those associated with more severe manifestations of a condition dominating (and eliminating the effect of) less serious ones.

***Principle 5*—The RXCs should not reward prescription proliferation**. The classification should not measure greater disease burden simply because more prescriptions are written. Hence, neither the number of times that a particular NDC appears, nor the presence of additional, closely related NDCs that indicate the same condition should increase predicted costs.

***Principle 6*—Providers should not be penalized for prescribing additional NDCs (monotonicity)**. This principle has two consequences for modeling: (1) no RXC should carry a negative payment weight, and (2) an RXC that is higher-ranked in a drug hierarchy (causing lower-rank drugs in the same hierarchy to be excluded) should have at least as large a payment weight as lower-ranked RXCs in the same hierarchy.

***Principle 7*—The classification system should be internally consistent (transitive)**. If category A is higher-ranked than category B in a hierarchy, and category B is higher-ranked than category C, then category A should be higher-ranked than category C. Transitivity improves the internal consistency of the classification system and ensures that the assignment of RXCs is independent of the order in which hierarchical exclusion rules are applied.

***Principle 8*—The classification should assign all NDCs (exhaustive classification)**. Because each diagnostic code potentially contains relevant clinical information, the classification should categorize all NDCs.

***Principle 9*—The classification should assign NDCs to only one RXC (mutually exclusive classification)**. Because each NDC can map to more than one RXC, the classification should map NDCs to the primary RXC based on route of administration, intended application of the product, ingredient list identifier, label, dosage form, strength of the drug, and other considerations.

***Principle 10*—Discretionary drug categories should be excluded from payment models**. RXCs that are particularly subject to intentional or unintentional discretionary prescribing variation or inappropriate prescribing by health plans or providers, or that are not clinically or empirically credible as cost predictors, should not increase cost predictions. Excluding these RXCs reduces the sensitivity of the model to prescribing variation, prescribing proliferation, and gaming.

In designing the RXCs, principles 6 (monotonicity), 7 (transitivity), 8 (exhaustive classification) and 9 (mutually exclusive classification), were generally followed. We used clinical and statistical assessments to make tradeoffs among other principles. For example, clinical meaningfulness (principle 1) is often best served by creating a very large number of detailed clinical groupings. However, a large number of groupings conflicts with adequate

sample sizes for each category (principle 3). We approached the inherent tradeoffs involved in designing a drug classification system using empirical evidence on frequencies and predictive power; clinical judgment on relatedness, specificity, and severity of RXCs; and professional judgment on incentives and likely provider responses to the classification system. The RXC risk adjustment model balances these competing goals to achieve prescription drug-based classes for use in risk adjustment.

A complete list of the RXCs is included as Appendix C.

4.2.5 Selecting Drug-Diagnosis Pairs for a Hybrid HHS-HCC Risk Adjustment Model

The development of a hybrid HHS-HCC risk adjustment model requires selecting drug-diagnosis pairs (RXC-HCC pairs) to include in the model,[42] from the more than 7,000 possible drug-diagnosis (RXC-HCC) pairs. Development is an iterative process that includes exploratory data analysis as well as recurring consultations with a panel of clinician consultants. Our work on this task is ongoing. In this paper, we have made a preliminary incomplete selection of pairs in order to illustrate a hybrid model (described in section 4.2.6). This section describes the criteria and methods we have used to make our initial drug-diagnosis pair selections. We solicit comments from the public on specific drug-diagnosis relationships that should be tested for incorporation into a hybrid HHS-HCC drug-diagnosis risk adjustment model, or methodologies to identify them.

Criteria for Selecting Drug-Diagnosis Pairs for the Hybrid Model

With the goal of gaining the advantages of drug information in the model while minimizing the disadvantages, we carefully considered of a wide range of issues. The main factors we used in making our selections were:

- We would seek to select drugs with patterns of non-discretionary prescribing.
- We would seek to avoid drugs where there are incentives for over-prescribing.
- We would seek to avoid drugs where there are variations in prescribing across providers, practices, and areas, which depends in part on whether prescription decisions are discretionary.
- We would carefully consider selection of high-cost drugs, as these costs may be the types of health risk variation across enrollee populations that risk adjustment is designed to account for, or if issuers know that risk adjustment transfers will compensate for the costs of these expensive drugs, then this compensation may reduce the incentives for issuers to strive for greater efficiency in prescription drug utilization.
- We would seek to avoid drugs indicated for multiple diagnoses.

[42] As discussed in section 4.2.3, an alternative approach is to focus on incremental statistical power alone of the RXC drug classes added to the HHS-HCC diagnosis-based model.

- We would seek to avoid drugs indicated for diagnoses not included in the HHS-HCC model.

- We would carefully consider selection of drugs in an area exhibiting a rapid rate of technological change to the extent possible, as a drug class that is associated with a specific, costly diagnosis in one year may no longer be commonly used for that condition the next, in which case the cost predictions based on previous years of data would be inaccurate.

All of these issues are given consideration in the design of the RXC classification system and selection of drug classes and their related HCCs for the draft hybrid models presented in this paper.

Empirical considerations

We performed a wide range of analyses on the 2014 MarketScan® adult sample (ages 21-64[43]) claims data — the data used to calibrate the HHS-HCC model — to select and refine the initial drug-diagnosis pairs for inclusion in the hybrid models. Early in this process, the two main conceptual approaches emerged to using drug information — diagnosis imputation or severity adjustment — and subsequent analyses were focused on finding drug-diagnosis pairs that would be well suited to these approaches. In this section, we describe the types of empirical relationships that we are considering in identifying drug classes to augment the HHS-HCC model, in identifying drug-diagnosis associations, and in investigating drug-diagnosis interactions.

Counts and mean expenditures by RXC. To gain a general idea of the composition of the RXC drug groupings that were created, we examined the following basic descriptive statistics for individuals with each of these drug indicators in their prescription claims: the total count of individuals in the sample who had a prescription drug claim in the RXC during the year, a count weighted by the number of months each person was enrolled, and mean expenditures for each RXC subsample (total expenditures, drug expenditures, and medical (non-drug) expenditures). These tabulations were useful to obtain a sense of how common each drug class is, and how the enrollees' expenditures vary across the RXCs. In particular, separating drug from non-drug expenditures provided some insight into whether the RXC predicts high expenditures because of associated medical treatments, or if those predicted costs were due primarily to the drugs themselves.

Drug-diagnosis associations (empirical data not shown). Because each drug class that enters the hybrid model must be assigned to a particular diagnosis (or group of diagnoses), we examined how closely-related each drug class is to each HHS-HCC. These analyses were conducted at the AHFS-TC (TC = "Rx therapeutic class") level as well as using the RXCs. We calculated three separate measures of association,[44] and created rankings of the most closely associated drug-diagnosis pairs. We used these lists to propose pairs that would otherwise have been overlooked, and those that seemed particularly closely linked were studied further in order

[43] The MarketScan data includes a small number of individuals age 65+.

[44] The three measures calculated were 1) odds ratio, 2) kappa statistic, and 3) positive predictive value.

to ascertain what information the drugs in the class could convey about the likely costs of people carrying the HCC diagnosis. This route of analysis was particularly helpful in discovering drug-diagnosis pairs that are well suited for an imputation relationship. AHFS-TC analyses were useful to determine whether the RXC groupings are appropriately implemented, or if any should be subdivided.

Stepwise regressions with drug classes and drug-diagnosis interactions. Stepwise models are a useful tool to examine which additional variables can add the most incremental predictive power to the existing HHS-HCC diagnosis-based models. In these analyses, we estimated a series of models, beginning with the baseline HHS-HCC model and adding one additional regressor in each round. Each added regressor was selected based on the criterion of achieving the maximum increase in the model's R- squared (R^2).

Adding drug classes to the current HHS-HCC model, either individually or collectively, results in modest, but not trivial, improvements in the model's overall predictive accuracy (R^2). However, most of the incremental predictive power from adding drugs to the HHS-HCC model is captured by a relatively small number of drug classes. In our hybrid model analyses (presented in section 4.2.6), we focus on the small number of drug classes that add the most predictive power.

These empirical exploratory analyses provide useful information about how to include drug information in the HHS-HCC model. The empirical considerations were supplemented with the clinical considerations discussed below.

Clinical Considerations

Clinical consultants, including doctors and pharmacists, provided deeper insights into the medical links between the health conditions and drug groups under consideration, and recommended specific interactions in addition to those produced by our empirical screens. The consultants were chosen to provide a wide range of up-to-date clinical expertise on treatments and protocols, and also because these individuals had familiarity with the principles and objectives of predictive risk modeling.

In these consultations we solicited suggestions for drug-diagnosis pairs to include in the models; obtained clarification on the clinical uses and indications of the drug classes being considered for inclusion; and reviewed lists of empirical results (stepwise regressions, etc.) to recommend additional drug-diagnosis pairs for consideration. Clinicians were thoroughly briefed on the different approaches to hybrid modeling, in particular the imputation and severity approaches. Discussions were conducted to achieve consensus-based decision-making.

Clinical considerations were particularly important in evaluating the "gameability" of each drug considered for inclusion in the model. Each candidate drug was discussed with the clinical consultants to assess how discretionary, variable, or "gameable" it is. Several drugs were rejected from the model specification based on this criterion. For example, drugs that are difficult to administer or carry severe side effects could be determined to be less "gameable." However, we do not assert that the drug classes chosen for analysis in an illustrative hybrid risk adjustment model (discussed in the next section) necessarily reflect drug classes for which physician prescribing is non-discretionary. For example, it is conceivable that physicians might

more quickly prescribe anti-diabetic drugs rather than treat diabetes with diet and exercise if risk adjustment established a financial incentive to do so.

Additional Considerations

Besides the criteria described above, there are additional model design considerations to consider. Two possible design elements that were investigated were as follows.

Imposing model restrictions based on days' supply or number of prescriptions. Because these models are intended to capture predictable cost variation – in particular the effects of chronic conditions on individual costs – we considered whether to require evidence of prolonged usage of a particular drug in order to trigger a drug indication. Prolonged usage could be signaled by the presence of multiple prescriptions of the same drug (or class of drugs), or by prescriptions totaling at least 30 or 60 days' supply. Our clinical consultants suggested a few RXCs for which a minimum days' supply restriction would be useful to distinguish severely ill patients from those with milder conditions. However, we determined that it was not appropriate to include these RXCs in the initial illustrative hybrid model specification presented in this paper. Nonetheless, these days' supply restrictions may be important to consider as we complete and refine the hybrid model for consideration — the public's input on whether days' supply restrictions should be imposed on drug classes added to the HHS-HCC model is solicited. We are particularly interested in feedback on which drug classes warrant days' supply requirements, and how many days' supply.

Subdividing/splitting RXCs, or including individual drugs. This question arose during discussions with the clinicians, in which the clinicians suggested that a specific drug or AHFS therapeutic class within an RXC is appropriately linked to an HCC, but other drugs in the RXC would confound this clinical connection. In each case, we considered whether to split an RXC or restrict an HCC-RXC interaction to certain drugs within the RXC. In these cases the drug-diagnosis relationship would have been more narrowly defined and thus have greater clinical precision. These advantages must be weighed against the added complexity, smaller sample size and less statistical stability, and the magnitude of the incremental predictive power. The public's input on particular drugs and drug-diagnosis combinations that should be incorporated into the HHS-HCC model is solicited.

Imputation only versus imputation-severity relationship. In theory, we could select some drug classes to use only in imputing diagnoses while other drug classes could be used to both impute a diagnosis and indicate its severity. However, in practice most or all drug classes that can be used to impute a diagnosis can also potentially be used to indicate the severity of a diagnosis. In the end, we do not find useful an *a priori* distinction between drug classes used for imputation only versus for both imputation and severity measurement. Instead, we only make an *a priori* distinction between drug classes used for (i) imputation and severity versus for (ii) severity only. For the imputation/severity classes of drugs, we examine different empirical models that include imputation-only versus imputation and severity (flexible) approaches.

Prophylactic use of drugs. One concern in imputing diagnoses from drug utilization is that drugs are sometimes used prophylactically in persons at risk of disease but who do not actually have the disease. For example, certain HIV antiretrovirals are used to prevent infection

before or after exposure to the virus. We know of no straightforward solution to the possibility of false imputations introduced by prophylactic usage of drugs.

Multiple indications for drugs. Another concern is that drug classes are often indicated for more than one diagnosis. For example, disease-modifying antirheumatic drugs (RXC 92.06, DMARDs) are most commonly used for rheumatoid arthritis (HCC 56), and less commonly for inflammatory bowel disease (HCC 48). Most people taking DMARDs have a rheumatoid arthritis (HCC 56) diagnosis, which would suggest the drug class can be used to impute missing HCC 56 diagnoses. However, some individuals take DMARDs for inflammatory bowel disease and do not have rheumatoid arthritis, hence it would be incorrect to always impute rheumatoid arthritis for users of DMARDs. To mitigate this issue, in our hybrid risk adjustment model (discussed in the next section), we impute rheumatoid arthritis (HCC 56) for people taking DMARDs (RXC 92.06) only if no diagnosis of inflammatory bowel disease (HCC 48) is present. However, for other drug classes indicated for multiple diagnoses where use of the drug is more evenly split among multiple diagnoses, adopting a similar approach may be more challenging.

4.2.6 Illustrative Hybrid Diagnosis-Drug Risk Adjustment Model Including Selected Diagnosis-Drug Pairs

In this section, we present summary results for initial versions of a hybrid drug-diagnosis model. Each version illustrates a different approach to incorporating drug information into the HHS-HCC risk adjustment models. We intend to further develop and refine these models over time. We request comments from the public on useful directions to expand, refine, or rework these illustrative models.

Drug-Diagnosis Pairs Included in the Hybrid Models

Table 4.4 shows the list of drug-diagnosis (RXC-HCC) pairs that are included in the initial hybrid models presented in this section. These pairs have been carefully chosen based on the criteria described above. Each pair is chosen as either an imputation/severity or a severity-only relationship. For each pair, the table shows how many people in the 2014 MarketScan® adult analytic file have the diagnosis (HCC), the drug utilization (RXC), and both. For RXC-HCC imputation-severity pairs, we can use these counts to determine how many people have the RXC drug flag but lack the HCC diagnosis, and can therefore be imputed to have the diagnosis. We also show these imputations expressed as a percentage of the HCC count. Lastly, we calculate the positive predictive value, which is defined as the proportion of people with the RXC who are also observed to have the HCC. This proportion is one measure of the strength of association between the RXC and the HCC.

Table 4.4:
Drug-Diagnosis (RXC-HCC) Pairs Chosen for the Initial Hybrid Risk Adjustment Models

					Counts			# of imputations	imputations as % of HCC count	positive predictive value
RXC	Label	HCC	HCC label	relationship	HCC	RXC	HCC and RXC			
8.03	HIV	001	HIV/AIDS	imputation/severity	29,247	36,274	25,497	10,777	0.368	0.703
8.11	Hep C Antivirals	037	Chronic Hepatitis	imputation/severity	26,722	6,218	3,268	2,950	0.110	0.526
24.01	Class IB and Class III Antiarrhythmics	142	Specified Heart Arrhythmias	imputation/severity	164,261	20,229	16,481	3,748	0.023	0.815
28.14	Antimanic agents	088	Major Depressive and Bipolar Disorders	imputation/severity	414,119	25,308	18,203	7,105	0.017	0.719
40.03	ESRD	184	End Stage Renal Disease	imputation/severity	12,840	8,941	6,761	2,180	0.170	0.756
48.05	Cystic Fibrosis tranmembrane conductance regulator agents	159	Cystic Fibrosis	imputation/severity	1,609	1,544	7	65	1,609	1,544
56.04	Anti-Inflammatory Agents Used to Treat Inflammatory Bowel Disease	048	Inflammatory Bowel Disease	imputation/severity	85,772	53,854	37,056	16,798	0.196	0.688
68.06 or	RXC group: Insulins and	019 OR	HCC Group: Diabetes	imputation/severity	1,056,797	1,022,463	815,060	207,403	0.196	0.797

					Counts			# of imputations	imputations as % of HCC count	positive predictive value
RXC	**Label**	**HCC**	**HCC label**	**relationship**	**HCC**	**RXC**	**HCC and RXC**			
68.07	Antidiabetics	020 OR 021								
92.05	Biologic Response Modifiers Acting on the Central Nervous System	118	Multiple Sclerosis	imputation/severity	39,414	25,666	23,357	2,309	0.059	0.910
92.06	Disease-modifying antirheumatic drugs (DMARDs)	056	Rheumatoid Arthritis and Specified Autoimmune Disorders	imputation/severity	134,683	71,864	45,293	26,571	0.197	0.630
24.06	High Severity Diuretic	130	Congestive Heart Failure	severity only	128,602	331,625	49,595	--	--	0.150
40.06	Ammonia Detoxicants	036	Cirrhosis of Liver	severity only	12,772	18,906	788	--	--	0.042
92.06	Disease-modifying antirheumatic drugs (DMARDs)	048	Inflammatory Bowel Disease	severity only	85,772	71,864	10,952	--	--	0.152

For example, consider the first pair listed, which matches RXC 8.03 (HIV) with HCC 1 (HIV/AIDS). In the MarketScan® sample, there are 29,247 individuals with the HCC diagnosis flag, and 36,274 individuals with the RXC drug utilization flag. Among these, 25,497 individuals in the sample have both the diagnosis (HCC) and a prescription for the drug class (the RXC flag). Based on these counts, we observed that there are 10,777 (= 36,274 – 25,497) individuals who have the drug but not the diagnosis; if we assume that everyone taking these HIV drugs has the associated HIV/AIDS condition, then this is the number of diagnoses we would impute based on the drug utilization. If we impute these diagnoses, the added individuals would represent a 36.8% increase in the number of individuals considered to have the HIV diagnosis above the number with the HCC in the sample. Lastly, the positive predictive value tells us that 70.3% (= 25,497/36,274) of individuals with the RXC indicator are also observed to have the HCC diagnosis in their records.

The drug-diagnosis pairs can include more than one RXC, more than one HCC, or both of these. For example, the list includes a diabetes drug-diagnosis relationship that includes two RXCs (insulins and antidiabetic drugs) as well as three HCCs (diabetes with acute complication, diabetes with chronic complication, and diabetes without complication). Either of the RXCs can be interpreted as an indication that the individual should have a diagnosis of one of these three diabetes HCCs. In addition, any RXC can be linked in the model to more than one HCC, and vice-versa. For example, RXC 92.06 (Disease-modifying antirheumatic drugs) acts as a severity indicator for HCC 048 (inflammatory bowel disease) and also acts as a severity indicator and an imputation indicator for HCC 056 (Rheumatoid arthritis and specified autoimmune disorders).

Hybrid Drug-Diagnosis Models

The remaining tables in this section present different versions of the initial hybrid model. Tables 4.5 and 4.6 show two different approaches to an imputation model; Table 4.7 shows results for the flexible specification that allows the data to determine the relationship between the drug and the diagnosis in each pair; and Table 4.8 shows a model focused on the drug-diagnosis pairs that are characterized with a severity only indicator relationship.

1) Imputation Only Model

The imputation-only model in Table 4.5 presumes that any individual with a particular health condition should be predicted to have the same incremental costs regardless of whether the condition is identified by a diagnosis, a drug prescription, or both. This is the most-constrained version of an imputation model, and it requires that the incremental expenditures predicted are equal for individuals with the HCC only, the RXC only, or both. This constraint is reflected in the fact that all three incremental cost columns in Table 4.5 are equal. We compare these estimates in Table 4.5 to a baseline HHS-HCC model that does not use drug information. For example, the predicted costs for someone with chronic hepatitis[45] in the baseline HHS-HCC model is $16,634. However, when we impute additional diagnoses for individuals who are

[45] Note that the HCC Chronic Hepatitis includes diagnoses other than chronic hepatitis C, such as chronic hepatitis B and non-viral chronic hepatitis.

taking a drug in RXC 8.11 (hepatitis C antivirals), the predicted incremental expenditures for this health condition are $25,425 regardless of whether the condition is identified by the drug indicator, the diagnosis, or both. In this case, the large increase in the size of the predicted expenditures compared to the baseline HHS-HCC model most likely reflects the extremely high costs of the hepatitis C drugs in this RXC. As a result, the additional individuals being included in this drug-diagnosis category are probably more costly than those who have the HCC diagnosis alone, and the hybrid model coefficient is therefore larger than the HCC model coefficient.

We see in Table 4.5 that this increase in the hybrid coefficient relative to the corresponding baseline HHS-HCC model coefficient is not universal; in many cases, the hybrid model coefficient is similar to the HCC model coefficient, and in other cases it is lower. For all HCCs in Table 4.5, however, additional individuals are imputed by usage of the drug class to have the HCC.

The R-squared for the baseline model and the imputation-model are virtually identical (0.3678 vs 0.3640). As stated, the imputation-only model is the most constrained version of the imputation model, which could be a reason why the R-squared is not higher than for the baseline. This is likely because the pure imputation-model is constraining individuals with the diagnosis (HCC) only (who are relatively cheap) to individuals taking the drug (who are relatively expensive). Note that the Rx dominant model is the same as the pure imputation-model, except it relaxes this constraint and as is shown below, the R-squared for the Rx dominant model is higher than in the baseline model.

Table 4.5:
Coefficient Estimates: Imputation-Only Hybrid Drug (RXC)-Diagnosis (HCC) Model

				Imputation-only Hybrid Model: incremental predicted costs			**Baseline HHS-HCC model**				
				R-squared = 0.3640			R-squared = 0.3678				
				Coefficients			**coefficient**	**count**	**Counts**		
RXC	**Label**	**HCC**	**HCC label**	**HCC only**	**RXC only**	**HCC and RXC**	**HCC**	**HCC**	**HCC only**	**RXC only**	**HCC and RXC**
8.03	HIV	001	HIV/AIDS	20,090	20,090	20,090	23,113	29,247	3,750	10,777	25,497
8.11	Hep C Antivirals	037	Chronic Hepatitis	25,425	25,425	25,425	16,634	26,722	23,454	2,950	3,268
24.01	Class IB and Class III Antiarrhythmics	142	Specified Heart Arrhythmias	10,442	10,442	10,442	10,056	164,261	147,780	3,748	16,481
28.14	Antimanic agents	088	Major Depressive and Bipolar Disorders	8,620	8,620	8,620	6,322	414,119	395,916	7,105	18,203
40.03	ESRD	184	End Stage Renal Disease	150,319	150,319	150,319	139,791	12,840	6,079	2,180	6,761
48.05	Cystic Fibrosis tranmembrane conductance regulator agents	159	Cystic Fibrosis	54,376	54,376	54,376	54,062	1,609	1,544	7	65
56.04	Anti-Inflammatory Agents Used to Treat Inflammatory Bowel Disease	048	Inflammatory Bowel Disease	12,318	12,318	12,318	12,956	85,772	48,716	16,798	37,056
68.06 or	RXC group: Insulins and	019 OR	HCC Group: Diabetes	4,556	4,556	4,556	4,826	1,056,797	241,737	207,403	815,060

				Imputation-only Hybrid Model: incremental predicted costs			Baseline HHS-HCC model				
				R-squared = 0.3640			R-squared = 0.3678				
				Coefficients			coefficient	count	Counts		
RXC	Label	HCC	HCC label	HCC only	RXC only	and RXC	HCC	HCC	HCC only	RXC only	and RXC
68.07	Antidiabetics	020 OR 021									
92.05	Biologic Response Modifiers Acting on the Central Nervous System	118	Multiple Sclerosis	40,047	40,047	40,047	38,978	39,414	16,057	2,309	23,357
92.06	Disease-modifying antirheumatic drugs (DMARDs)	056	Rheumatoid Arthritis and Specified Autoimmune Disorders	27,387	27,387	27,387	15,101	134,683	89,390	26,571	45,293

The four rightmost columns on this table show the counts for each way in which individuals can be identified with a health condition. In order, these show 1) the total number of people in the sample with the HCC flag; 2) the number of people with the HCC flag but no drug flag; 3) the number of people with the RXC flag but no HCC; and 4) the number of people with both the RXC and the HCC. The "RXC only" column indicates the number of people imputed to have a diagnosis based on the presence of a drug class. For example, the number of people imputed by hepatitis C anti-viral drug usage to have chronic hepatitis is 2,950, which are added to the 26,722 people with a chronic hepatitis diagnosis to estimate the model.

2) Rx Dominant Model

Table 4.6 presents results from a less restrictive version (than imputation only) of the hybrid model, which we refer to as the "Rx dominant" or "drug-dominant" model. In this version, we estimate a predicted incremental expenditure for people who are identified by their HCC flag, and estimate a different predicted expenditure for people who are identified by their RXC flag. This model restricts people with both the HCC and the RXC to have the same incremental predicted cost as those with the RXC alone. In this way, the RXC (drug use) predicts greater incremental costs than for someone with only a diagnosis, which allows the drug information to convey some degree of illness severity in addition to its primary role imputing a (presumably) missing diagnosis. This is why we call this version of the model "drug-dominant," because the cost prediction is the same when the drug use is present, whether or not the diagnosis is present.

Once again, the chronic hepatitis drug-diagnosis pair provides a useful illustration of how this model design works. Individuals with only the chronic hepatitis HCC 37 (and not the hepatitis C drug class) are predicted to have an additional $2,436 of expenditures (in addition to their age/sex cell coefficient and any other HCCs they have). Individuals flagged with RXC 8.11 (hepatitis C antivirals), however, are predicted to generate $109,789 of incremental expenditures. This large gap is primarily due to the high costs of recent drugs to treat hepatitis C. Individuals who have both the HCC and the RXC in their claims records are restricted to have the same incremental costs as someone with the RXC alone. We also examined the number of individuals in the sample that were expected to reflect these cost estimates. Based on the model, we observed that 23,454 people have only the HCC, and will therefore be predicted to incur the $2,436 of incremental expenditures. A total of 6,218 people in the sample have a flag for the RXC, and thus the higher incremental predicted costs of $109,789; these are comprised of 3,268 who also have the HCC diagnosis, and 2,950 who do not. This example also illustrates how inclusion of RXCs can potentially influence plan design incentives in ways that could lead to higher utilization of prescription drugs – an increase of this amount could weaken the incentive for plans to engage in medical management that limits access to a high cost drug only to individuals with clinical need.

The R-squared for the Rx dominant model (0.3845) is higher than for the baseline model (0.3678) by 1.7 percentage points. Thus, while the pure imputation-model R-squared was virtually the same as for the baseline model, R-squared for the Rx dominant model was slightly higher. As suggested, this is likely due to the Rx dominant model predicting higher expenditures for individuals utilizing expensive drug classes, i.e., not constraining predicted expenditures for

people with only the diagnosis (HCC) (who are relatively inexpensive) to predicted expenditures for people taking the drug (who are relatively expensive), as is done in the pure imputation model.

In most of these ten drug-diagnosis pairs, there is a large difference between the estimated costs of individuals identified by their prescription drug utilization and those who are identified only from a diagnosis. This suggests that, for these pairs, the drug-dominant specification captures important distinctions in the expenditures depending on whether the health condition is recorded by a diagnosis or by prescription drug information, with the latter often indicating utilization of an expensive drug class.

3) Flexible Hybrid Model

The third hybrid model is the "flexible hybrid," a terminology which reflects the model's allowance for three different predicted levels of incremental expenditure: one for people with the diagnosis only, one for those with prescription drug claim only, and a third level for people with both indicators. Table 4.7 presents results for the flexible hybrid model. Consider HCC 142, specified heart arrhythmias, as an example. Predicted incremental costs for an individual with only the diagnosis are $8,544. People flagged with only RXC 24.01 (Class IB and Class III Antiarrhythmics), however, are expected to incur $21,551 of expenditures for this health condition, and the presence of both the RXC and the HCC are associated with costs of $25,727. Note that in the absence of any drug utilization information, 3,748 people (RXC only column) in the sample would not have their risk scores reflect the likelihood that they will incur any of these extra costs. In this hybrid model, however, those individuals would have additional predicted costs of $21,551 based on the information contained in their drug claims. This illustrates the types of improvements in diagnosis identification that may be obtained from incorporating drug utilization in risk adjustment. As with the Rx dominant model, the flexible hybrid model conveys some degree of illness severity in addition to its primary role imputing a (presumably) missing diagnosis. In addition, the R-squared for the flexible hybrid model is almost two percentage points higher than for the baseline model (0.3861 vs 0.3678), similar to that of the Rx dominant model.

One potential problem with this unrestricted specification, however, is the possibility that a risk score could actually be reduced by the presence of a diagnosis in the individual's claims. We see this occurring in two of the drug-diagnosis pairs: chronic hepatitis (HCC 37) and multiple sclerosis (HCC 118). In multiple sclerosis, for example, the model estimates incremental expenditures of $57,020 for individuals who are prescribed a drug in RXC 92.05, biologic response modifiers. Predicted expenditures for individuals with the HCC as well as the RXC, however, are lower: $54,993. This implies a risk score *reduction* for adding an HCC 118 diagnosis to anyone who is prescribed one of these drugs. This violates one of the basic principles of risk adjustment modeling, which is that there should be no penalty for recording additional (accurate) diagnoses. Before implementing any hybrid model, these types of perverse reporting incentives should be minimized in the model design and, if necessary, eliminated by imposing appropriate constraints on the affected coefficient estimates.

4) Severity-Only Hybrid Model

The final set of hybrid model results shown, in Table 4.8, constrain the model to a "severity only" form; that is, the presence of a prescription in the drug class signals a more severe case of the related diagnosis, and thus is likely to incur greater medical expenditures relative to someone without the drug, but does not impute the diagnosis in cases where diagnostic information is not available. Unlike the imputation approach, we do not need the drug to be a reliable indicator that someone is affected by the health condition. For example, most people taking disease-modifying antirheumatic agents (DMARDs, RXC 92.06) do not have inflammatory bowel disease (HCC 48). People with HCC 48 who do take DMARDs, however, are likely to have more severe cases of this condition than those who are not taking the drug. This relationship is what the severity model is designed to measure.

Table 4.8 is the only model in which we include the three severity only drug-diagnosis pairs from Table 4.4:

- High-severity diuretics (RXC)/congestive heart failure (HCC);
- Ammonia detoxicants (RXC)/cirrhosis of liver;
- DMARDs (RXC)/inflammatory bowel disease (HCC).

These pairs are not included in the previous tables because the drug class does not necessarily indicate the presence of the diagnosis, i.e., it does not reliably impute the diagnosis (because these drug classes are used for other diagnoses).

The drug-diagnosis pair, DMARDs with inflammatory bowel disease, provides a clear illustration of what the model estimates indicate. In the severity approach, there are two different levels of predicted incremental expenditures: a lower level for individuals flagged with the HCC only ($9,579), and a higher level for those flagged with the HCC and the RXC ($35,885). People who have the drug but not HCC 48 do not receive "credit" for any additional expenses, because the severity model assumes the drug utilization conveys information about severity only when a specific diagnosis is present. Most people in the sample have the HCC only, but for the relative few (10,952) who have the RXC as well, incremental expenditures are quite a bit larger. Compare this pair of predictions with the single $12,956 estimate for all HCC 48 individuals in the baseline HHS-HCC model. This illustrates how predictive accuracy can be improved with the severity approach. Finally, the R-squared for the severity only model is close to one percentage point higher than for the baseline model (0.3755 vs 0.3678).

Table 4.6:
Coefficient Estimates: Rx-Dominant Hybrid

				Rx-Dominant Hybrid Model: incremental predicted costs			Baseline HHS-HCC model				
				R-squared = 0.3845			R-squared = 0.3678				
				Coefficients			coefficient	count	Counts		
RXC	Label	HCC	HCC label	HCC only	RXC only	HCC and RXC	HCC	HCC	HCC only	RXC only	HCC and RXC
8.03	HIV	001	HIV/AIDS	1,862	22,879	22,879	23,113	29,247	3,750	10,777	25,497
8.11	Hep C Antivirals	037	Chronic Hepatitis	2,436	109,789	109,789	16,634	26,722	23,454	2,950	3,268
24.01	Class IB and Class III Antiarrhythmics	142	Specified Heart Arrhythmias	8,554	25,056	25,056	10,056	164,261	147,780	3,748	16,481
28.14	Antimanic agents	088	Major Depressive and Bipolar Disorders	6,111	8,985	8,985	6,322	414,119	395,916	7,105	18,203
40.03	ESRD	184	End Stage Renal Disease	101,862	152,554	152,554	139,791	12,840	6,079	2,180	6,761
48.05	Cystic Fibrosis tranmembrane conductance regulator agents	159	Cystic Fibrosis	45,087	272,137	272,137	54,062	1,609	1,544	7	65
56.04	Anti-Inflammatory Agents Used to Treat Inflammatory Bowel Disease	048	Inflammatory Bowel Disease	9,276	9,340	9,340	12,956	85,772	48,716	16,798	37,056
68.06 or	RXC group: Insulins and	019 OR	HCC Group: Diabetes	2,792	4,757	4,757	4,826	1,056,797	241,737	207,403	815,060

				Rx-Dominant Hybrid Model: incremental predicted costs			Baseline HHS-HCC model				
				R-squared = 0.3845			R-squared = 0.3678				
				Coefficients			coefficient	count	Counts		
RXC	Label	HCC	HCC label	HCC only	RXC only	HCC and RXC	HCC	HCC	HCC only	RXC only	HCC and RXC
68.07	Antidiabetics	020 OR 021									
92.05	Biologic Response Modifiers Acting on the Central Nervous System	118	Multiple Sclerosis	15,177	55,136	55,136	38,978	39,414	16,057	2,309	23,357
92.06	Disease-modifying antirheumatic drugs (DMARDs)	056	Rheumatoid Arthritis and Specified Autoimmune Disorders	8,659	27,799	27,799	15,101	134,683	89,390	26,571	45,293

				Flexible Hybrid Model: incremental predicted costs			Baseline HHS-HCC model				
				R-squared = 0.3861			R-squared = 0.3678				
				Coefficients			coefficient	count	Counts		
RXC	Label	HCC	HCC label	HCC only	RXC only	HCC and RXC	HCC	HCC	HCC only	RXC only	HCC and RXC
8.03	HIV	001	HIV/AIDS	1,849	15,060	25,869	23,113	29,247	3,750	10,777	25,497
8.11	Hep C Antivirals	037	Chronic Hepatitis	2,849	115,448	104,901	16,634	26,722	23,454	2,950	3,268
24.01	Class IB and Class III Antiarrhythmics	142	Specified Heart Arrhythmias	8,544	21,551	25,727	10,056	164,261	147,780	3,748	16,481
28.14	Antimanic agents	088	Major Depressive and Bipolar Disorders	6,114	5,135	10,390	6,322	414,119	395,916	7,105	18,203
40.03	ESRD	184	End Stage Renal Disease	101,879	89,226	172,249	139,791	12,840	6,079	2,180	6,761
48.05	Cystic Fibrosis tranmembrane conductance regulator agents	159	Cystic Fibrosis	45,076	199,091	278,920	54,062	1,609	1,544	7	65
56.04	Anti-Inflammatory Agents Used to Treat Inflammatory Bowel Disease	048	Inflammatory Bowel Disease	9,365	6,440	10,624	12,956	85,772	48,716	16,798	37,056
68.06 or	RXC group: Insulins and	019 OR	HCC Group:	2,767	2,001	5,401	4,826	1,056,797	241,737	207,403	815,060

				Flexible Hybrid Model: incremental predicted costs			Baseline HHS-HCC model				
				R-squared = 0.3861			R-squared = 0.3678				
				Coefficients			coefficient	count	Counts		
68.07	Antidiabetics	020 OR 021	Diabetes								
92.05	Biologic Response Modifiers Acting on the Central Nervous System	118	Multiple Sclerosis	15,184	57,020	54,993	38,978	39,414	16,057	2,309	23,357
92.06	Disease-modifying antirheumatic drugs (DMARDs)	056	Rheumatoid Arthritis and Specified Autoimmune Disorders	8,649	27,147	28,065	15,101	134,683	89,390	26,571	45,293

Table 4.8:
Coefficient Estimates: Severity-Only Hybrid Model

				Severity Hybrid Model: R-squared = 0.3755 Coefficients			Baseline HHS-HCC model R-squared = 0.3678 coefficient	count	Counts		
RXC	Label	HCC	HCC label	HCC only	RXC only	HCC and RXC	HCC	HCC	HCC only	RXC only	HCC and RXC
8.11	Hep C Antivirals	037	Chronic Hepatitis	3,981	0	104,956	16,634	26,722	23,454	*	3,268
24.01	Class IB and Class III Antiarrhythmics	142	Specified Heart Arrhythmias	8,398	0	25,254	10,056	164,261	147,780	*	16,481
24.06	High Severity Diuretic	130	Congestive Heart Failure	9,144	0	14,682	11,597	128,602	79,007	*	49,595
28.14	Antimanic agents	088	Major Depressive and Bipolar Disorders	6,133	0	10,448	6,322	414,119	395,916	*	18,203
40.06	Ammonia Detoxicants	036	Cirrhosis of Liver	15,710	0	19,336	15,874	12,772	11,984	*	788
68.06 OR 68.07	RXC group: insulins/antidiabetics	019 OR 020 OR 021	HCC Group: Diabetes	2,936	0	4,889	4,826	1,056,797	241,737	*	815,060
92.05	Biologic Response Modifiers Acting on the Central Nervous System	118	Multiple Sclerosis	15,114	0	54,936	38,978	39,414	16,057	*	23,357
92.06	Disease-modifying antirheumatic drugs (DMARDs)	048	Inflammatory Bowel Disease	9,579	0	35,885	12,956	85,772	74,820	*	10,952
92.06	Disease-modifying antirheumatic drugs (DMARDs)	056	Rheumatoid Arthritis and Specified Autoimmune Disorders	8,602	0	27,412	15,101	134,683	89,390	*	45,293

Table 4.7:
Coefficient Estimates: Flexible Hybrid

4.2.7 Initial Evaluation of Alternative Hybrid Drug-Diagnosis Risk Adjustment Models

In this section we offer some brief comments on the illustrative hybrid models presented in section 4.2.6, with reference to the evaluation criteria listed in section 4.2.2. Development of the hybrid models is ongoing, and our assessment of them is preliminary, incomplete, and subject to change.

Criterion 1: Clinical/Face Validity

The drug-diagnosis pairs used in all of the models were required to have clinical face validity as part of the process of selecting these pairs. However, the models that utilize drug usage to indicate severity of diagnosis (i.e., models other than the pure imputation model) probably have a greater degree of clinical face validity. This is because use of a drug class will typically contain some information about severity of illness. The Rx dominant model may also have greater clinical face validity than the completely flexible model specification, because it may not be clear why the presence of a reported diagnosis should affect incremental predicted cost when the drug class imputes the diagnosis.

Criterion 2: Empirical/Predictive Accuracy

The hybrid models we estimated have broadly similar overall predictive accuracy (R-squared). But the models imposing fewer restrictions on the estimated relationship between drugs and diagnoses—the flexible hybrid model in particular—are the most data-driven and fit the observed data the most closely. The most highly-constrained models—the severity only model and the imputation model—have the least predictive accuracy. In addition to overall predictive accuracy, predictive accuracy for groups is important. For the groups that the hybrid models are adjusting for (e.g., individuals utilizing hepatitis C antivirals in an Rx dominant model), there will be substantial gains in predictive accuracy. The hybrid models that add the most predictive accuracy are those that predict higher expenditures for individuals utilizing expensive drug classes. On the one hand, this may promote access to these drug classes; on the other hand, discretionary utilization of these drug classes will be rewarded and cost control incentives weakened.

Criterion 3: Incentives for Prescription Drug Utilization

All of the hybrid models create incentives for health plans' contracted providers to prescribe the drug classes used in the model. The imputation only and severity only models create the least strong incentives. The imputation model predicts the same incremental costs if a condition is identified through a reported diagnosis, drug utilization, or both. If a diagnosis is reported, there is no incremental predicted cost associated with the associated drug class prescription. The imputation only model does provide incentives for the providers to potentially overprescribe where they may not have otherwise. The severity only model predicts no incremental cost if the diagnosis is not present, but may create some incentives to prescribe drug treatments because the treatment will increase the incremental cost if a specific diagnosis is also present. The Rx dominant and flexible models create the strongest incentives for drug

prescribing, because incremental predicted cost increases the most with drug utilization in these models.

Criterion 4: Sensitivity to Variations in Prescription Drug Utilization

All of the hybrid models are sensitive to variations in prescription drug utilization. The same models that create the weakest and strongest incentives for drug utilization are the least and most sensitive to variations in drug utilization.

Criterion 5: Incentives for Diagnosis Reporting

When drug utilization is present, the incremental predicted cost from neither the imputation only nor the Rx dominant models are affected by the reporting of associated diagnoses. The flexible model is sensitive to the reporting of associated diagnoses even when drug use is present, and as discussed in section 4.2.6, the flexible model can lead to lower predicted incremental cost when a diagnosis is reported than when it is not reported. The severity only model is sensitive to the reporting of diagnoses, because the higher cost associated with a drug-diagnosis pair is only recognized when the diagnosis is reported.

4.2.8 Discussion

Based on the research performed so far, we believe that a hybrid model that includes prescription drug data in the HHS-HCC risk adjustment framework deserves consideration. This revision would need to be carefully designed and implemented, with an understanding that the potential gains in predictive power, accuracy, and fairness will come with costs of increased potential for gaming, incentives for greater prescription drug utilization, sensitivity of risk adjustment to variations in drug utilization unrelated to enrollee health status, and added administrative burden. We solicit comments from the public on the desirability of including prescription drugs in risk adjustment and an appropriate approach for doing so.

Gains in accuracy and fairness include the ability to impute diagnoses for enrollees in plans that may not completely report diagnoses, as well as being able to more precisely identify severely-ill individuals separately from those with milder cases of the same health conditions. These refinements are worth evaluating, though as we point out, it is difficult to identify specific drug utilization indicators that can unambiguously improve model performance without raising new concerns about incentives and fairness. Moreover, the gains in model predictive accuracy from incorporating prescription drug utilization appear to be quite modest, and arise from identifying individuals utilizing expensive drugs.

Before moving ahead with a hybrid risk adjustment model, CMS would first assess several factors, including (i) the operational costs, both for issuers as well as for the department overseeing data operations; (ii) what drug classes, or drug-diagnosis pairs, should be incorporated into the model; and (iii) whether the model should take an imputation approach, a severity approach, or a combination of both. For example, starting with a relatively limited number of drug classes in the model specification seems advisable.

Lastly, all the research discussed in this paper has been conducted with the HHS risk adjustment adult models and sample; before proceeding, we would need to verify whether drug

information can improve the performance of the child model as well, and also consider whether there is a role for drug utilization measures in the infant model.

4.3 High Risk Enrollee Pooling in HHS Risk Adjustment

Traditional risk adjustment does not predict the presence of extremely high-cost enrollees with precision since predicted plan liabilities reflect the average cost for individuals with a given set of age, sex, and diagnosis characteristics. Since the distribution of spending is skewed toward high-cost enrollees, these individuals are responsible for a large portion of total spending. As a consequence, even with risk adjustment in place, issuers retain an incentive to engage in risk selection in order to avoid these very high-cost enrollees. To mitigate any such residual incentive for risk selection, we could seek to insulate issuers against these high cost enrollees through the risk adjustment methodology.

In greater detail, risk adjustment uses a model to predict costs based on certain identifying characteristics of a population. In the HHS-HCC model, age, sex, health conditions (as established by certain diagnoses), metal level, and cost-sharing reduction status are used to predict costs. This modeling process works well with average people in each category of characteristics. For each category of characteristics, the model underpredicts the costs of people whose costs are far above the average (such as those with expensive, acute events during the risk adjustment year), and similarly, overpredicts the costs of people whose costs are far below the average. Although risk adjustment seeks to accurately predict the average health risk for the entirety of a health plan's enrollees, to the degree that the plan experiences extreme cost outliers, this goal might not be realized. The potential problem of extreme cost outliers affects the process of establishing the risk adjustment factors (through calibration) as well as the predictive power of the model in risk adjustment operations.

To give an example of the effect of outliers on the model calibration, assume there are 1,000 individuals with condition X. In general, the costs of condition X average $50,000. In this same population of individuals with condition X, there is one individual whose costs were $2 million. In calibration, this person would change the predicted average costs in the model by the value of about $2,000,000/1,000 individuals, or approximately an additional $2,000. So now the model predicts the average expenditures for condition X as approximately $52,000 as opposed to $50,000. In operation, when the model is applied, every enrollee with condition X would receive the average value of $52,000.[46] For the typical enrollee with condition X, the issuers would be overpaid by $2,000, while the issuer with the enrollee with a cost of $2 million would not receive any extra compensation.

Such a result would have two potentially undesirable consequences. First, it means that risk adjustment would not fully eliminate issuers' incentives to engage in risk selection. Because risk adjustment does not fully compensate for the additional costs associated with high-cost individuals in any given observed category, issuers may benefit from taking steps to avoid the high-cost individuals in any particular category, with negative consequences for consumers.

[46] In practice, risk adjustment works on relative values rather than actual dollars; however, in order to demonstrate this example we will use actual dollars.

Second, to the extent that commercial reinsurance cannot adequately limit the risk of these enrollees, this scenario may leave issuers, particularly smaller issuers, vulnerable to unpredictable costs, which could have implications for the stability of the individual and small group markets.

We are considering whether to address this problem by reducing issuers' exposure to outliers via modifications to the HHS risk adjustment model. First, when calibrating risk adjustment, we would modify the calculation of enrollee-level plan liability to exclude a percentage of costs above a certain threshold level. The existing risk adjustment transfer formula would then be applied to the modified plan liability risk scores in the current fashion. Then, to account for the costs associated with the high cost outliers, we would calculate an additional transfer amount for each issuer. This amount would be determined by calculating the specified percentage of the costs above the threshold for each of an issuer's high cost enrollees and then summing these amounts across all issuers. This total amount spread across all issuers would then determine a charge or adjustment to all issuers' transfer amounts.

In considering this change to the model, we are considering a uniform adjustment across State markets across the country, since such a pooling would be most effective in reducing the impact of extreme high cost outliers. We recognize that creating a uniform pool of high cost enrollees, by risk pool or market, could result in some States or geographic areas subsidizing issuers with high-cost enrollees in other States or markets. We note that while this adjustment would occur in a uniform manner across States and markets, we would continue to calculate risk adjustment risk scores and transfers using the recalibrated, truncated model in a risk pool in a market in a State.

The lower the high-cost threshold used to implement this approach, the better the predictive power of the HHS risk adjustment model. Additionally, the lower the thresholds, the more enrollees (and costs) that will be in a uniform pool. While more issuers would be eligible to receive uniform pool payments with a lower threshold, the charge or adjustment to fund the pool would also be higher for all issuers, the lower the threshold. We are considering using a threshold of approximately $1 million, pending further analysis. At this threshold level, we expect net transfers across States to be generally small as a percent of premiums.

The other consideration is the percent of costs to be reimbursed from the uniform pool above the threshold. With a greater percentage, the predictive power of the model improves and compensation of the actuarial risk of high cost enrollees becomes more complete. If 100 percent of costs above the threshold were reimbursed, the model would have greatest predictive power and payments would be most complete. However, reimbursing 100 percent of costs would eliminate any incentives to control costs on the part of the issuer for any individual whose costs exceed the threshold value.

To provide an illustration of how this would work, assume a threshold value of $1 million and reimbursement of 90 percent of costs above the threshold. Using the same condition X example as above, an outlier enrollee with costs of $2 million exceeds the threshold of $1 million by $1 million. By removing 90 percent of costs above the threshold, the modified model removes all but $100,000 from the enrollee's costs above $1 million, making the costs in calibration $1 million + $100,000, totaling $1,100,000. In this case, still assuming 1,000 individuals with condition X with an average cost of $50,000, the outlier enrollee would add

about $1,100 ($1,100,000/1,000) into the cost of condition X, meaning the model's overprediction is reduced from the previous approximately additional $2,000 value. This adjustment alone improves the predictive power of the model for the typical enrollee.

In operation, the costs that fall above the thresholds would be placed into the uniform risk adjustment high cost pool. In this case, the costs that were removed are $2 million – $1,100,000 = $900,000. The issuer who has this enrollee would be paid the approximately $51,100 predicted in the model for condition X, plus the $900,000 above the threshold, or about $951,100. We are continuing to evaluate the parameters and impacts of this proposal. We welcome feedback on implementing this approach, including how to do so in a budget neutral manner, and whether we should attempt to do so for the 2018 or 2019 benefit year recalibration.

4.4 Use of a Concurrent Model

Risk adjustment models can only utilize available information to predict expenditures. As discussed in Chapter 1, most risk adjustment models used for payment are "prospective," meaning they use prior year information to predict current year medical expenditures. Prospective models tend to be favored because they tend to tend to emphasize the impact of ongoing chronic conditions on costs (as opposed to random current year costs that can be pooled as "insurance risk"). A concurrent model uses current year information to predict current year costs. Concurrent models tend to emphasize the prediction of costs associated with current year acute health events. The HHS risk adjustment model is concurrent.

The predictive power of concurrent models over prospective models is consistent with previous findings where much of the improvement comes from unpredictable costs such as cancer diagnoses, strokes, heart attacks, and other event-related costs that occur during the year for which costs are being predicted. Because concurrent risk adjustment explains more of the variation in current (acute) costs, it reduces unsystematic risk, which may benefit small health plans that do not have enough enrollees to diversify away unsystematic risk. The prospective models, on the other hand, are effectively trying to predict, on the basis of prior encounters for conditions, the individuals who will experience such events in the coming year. The improvement in fit for concurrent modeling arises in large part from capturing the costs of treating emerging acute conditions or exacerbations of chronic conditions.

When the Affordable Care Act was first established, a prospective model was infeasible due to the lack of previous year information on health status (diagnoses), and also the fact that unlike Medicare, enrollees move in and out of enrollment in the individual and small group markets, so prior year diagnostic data was not available for all enrollees. We have received feedback that HHS should consider a prospective model. However, our use of a concurrent model supports the intent of the Affordable Care Act – encouraging choice, competition, and growth in plans. Because a prospective model cannot easily reflect enrollees' movement between markets and across issuers, we believe a concurrent model is more appropriate for the individual and small group markets.

4.5 Data for 2018 Recalibration

We have used the three most recent years of MarketScan® data to recalibrate the 2016 and 2017 benefit year risk adjustment models. This approach has allowed for using the blended,

or averaged, coefficients from three years of separately solved models, which promotes stability for the risk adjustment coefficients year over year, particularly for conditions with small sample sizes. This approach in previous years has also required that we finalize coefficients based on data that does not become available until after the publication of the proposed Payment Notice.

We received several comments to the 2017 Payment Notice proposed rule requesting that the Payment Notice schedule be moved up to accommodate substantive comments. We also received many comments on how to best address the data lag for HHS risk adjustment and better reflect new treatments that may be associated with high cost conditions. As we noted above, we took steps in the 2017 Payment Notice to more accurately trend specialty and traditional drug growth separately from medical and surgical expenditures. However, with the blended, three year data coefficients, any introductions of new costs for particular conditions are still weighted by two years of older data. We are exploring changing the tradeoff between stability and data recency by reweighting the different years' data, including perhaps using only the latest year's data. For example, for 2018 risk adjustment, the most recent dataset would be 2014 MarketScan®. We could recalibrate the 2018 risk adjustment models solely on 2014 MarketScan®, similar to the original calibration on 2010 MarketScan® data. We believe 2014 MarketScan® reflects the introduction of several notably expensive treatments associated with high cost conditions and should be reflective of changes in treatment patterns, to the extent possible. We seek comments on this approach.

4.6 Data for 2019 Recalibration

4.6.1 Current EDGE Data Environment

CMS developed an approach to perform and evaluate calculations for both the risk adjustment program and the transitional reinsurance program.[47] In evaluating data collection options, CMS determined that a distributed data collection model proved the most effective approach for obtaining and processing the data necessary for both reinsurance and risk adjustment calculations because such a model would ensure minimal transfer of protected health information between issuers and CMS, thereby lowering privacy and data security risks; and standardization of business processes, timing and rules.

We implemented a distributed data approach through External Data Gathering Environment or "EDGE" servers. Issuers in States where HHS is operating the risk adjustment or reinsurance program upload enrollee, pharmaceutical claim, medical claim, and supplemental diagnosis information from their proprietary systems to an issuer-owned and controlled EDGE server. (Issuers have the option to own and operate the server themselves, or to have a third-party entity operate the server.) The EDGE server runs CMS-developed software designed to verify submitted data and execute the risk adjustment and reinsurance processes. Detailed data, file processing metrics, and outbound data files are provided to issuers, while only plan

[47] Section 1341 of the Affordable Care Act established the reinsurance program as a temporary three-year program that provides payments to issuers of non-grandfathered individual market plans both inside and outside of the Marketplaces for the 2014-2016 benefit years. Reinsurance payments are based on a coinsurance rate or proportion of an issuer's claims for an individual enrollee costs that are above an attachment point and below a reinsurance cap for the applicable benefit year.

summarized data and file processing metrics are provided to CMS. CMS does not receive any individual-level data as part of this process.

The current distributed data model meets the objectives outlined above; however, the lack of enrollee-level data also limits our ability to use that data to improve the risk adjustment program. Specifically, with access to de-identified enrollee-level data, we could recalibrate the risk adjustment models beginning in the 2019 benefit year, potentially using more recent 2016 benefit year data from the actual population to which risk adjustment applies, instead of using commercial databases that approximate the individual and small group market populations. To that end, we are considering modifying the current distributed data collection approach so that relevant enrollee-level EDGE data from individual and small group enrollees can be incorporated into recalibration, while maintaining the underlying goals of the distributed data approach.

We note that we may also be able to use this or other enrollee-level data to recalibrate the model to also account for socioeconomic status. For example, we could do so by using cost sharing reductions (CSRs) as a proxy for that status. We could also do so through other approaches, but would need to implement those approaches in such a manner as to avoid the transmission of potentially identifying data. For example, with information on an enrollee's zip code, it is possible to determine a wide variety of socioeconomic characteristics of the neighborhood in which the enrollee lives using published estimates from the Census Bureau's American Community Survey, including median income and average educational attainment. These characteristics could then be used to calibrate the risk adjustment model. Note that implementing this approach would not necessarily require the transmission of identifying enrollee-level information such as zip code. Rather, we could have the CMS EDGE server software match the zip code to the applicable risk adjustment calibration variable (for example, through a table matching zip codes to neighborhood median income), and have the de-identified variable transmitted as part of the data set, not the zip code.

The following sections discuss a method for importing the enrollee-level data issuers already submit to their EDGE server to CMS without identifying any individual issuer and while maintaining the privacy and security of the data.

4.6.2 Variables Needed for Recalibration Using EDGE Risk Adjustment Data

EDGE server data could be used to recalibrate the risk adjustment model with enrollee and claims level in a format that aligns with the components of the risk adjustment model, if CMS is able to access select data elements in the enrollee, medical claim, pharmacy claim and supplemental diagnosis files for the EDGE servers. CMS would not have access to geographic rating area, State, or enrollee, plan, or issuer identifiers for any of the potential enrollee-level data. The EDGE data elements we are considering are identified by EDGE server submission file type in Table 4.9.

Table 4.9:
Data Elements Required for Enrollee-Level EDGE Data Collection

Data Element	EDGE File Type
Masked Enrollee ID	Enrollee
Enrollee Birth Date	Enrollee

Data Element	EDGE File Type
Enrollee Gender	Enrollee
Subscriber Indicator	Enrollee
Masked Subscriber ID	Enrollee
Metal Level Identifier	Enrollee
Enrollment Start Date	Enrollee
Enrollment End Date	Enrollee
Enrollment Period Activity Indicator	Enrollee
Policy Premium Amount	Enrollee
Masked Enrollee ID	Medical Claim
Form Type (institutional or professional)	Medical Claim
Claim Identifier	Medical Claim
Original Claim Identifier	Medical Claim
Claim Processed Date and Time	Medical Claim
Bill Type Code	Medical Claim
Void Replace Code	Medical Claim
Diagnosis Type Code	Medical Claim
Diagnosis Code	Medical Claim
Discharge Status Code	Medical Claim
Statement Cover From Date	Medical Claim
Statement Cover To Date	Medical Claim
Billing Provider ID Qualifier	Medical Claim
Billing Provider Identifier	Medical Claim
Issuer Claim Paid Date	Medical Claim
Allowed Total Amount	Medical Claim
Policy Paid Total Amount	Medical Claim
Derived Service Claim Indicator	Medical Claim
Service Line Number	Medical Claim Service Line
Service From Date	Medical Claim Service Line
Service To Date	Medical Claim Service Line
Revenue Code	Medical Claim Service Line
Service Type Code	Medical Claim Service Line
Service Code	Medical Claim Service Line
Service Modifier Code	Medical Claim Service Line
Service Facility Type Code	Medical Claim Service Line
Rendering Provider ID Qualifier	Medical Claim
Rendering Provider Identifier	Medical Claim
Allowed Amount	Medical Claim Service Line
Policy Paid Amount	Medical Claim Service Line
Derived Service Claim Indicator	Medical Claim Service Line
Masked Enrollee ID	Pharmacy

Data Element	EDGE File Type
Claim Identifier	Pharmacy
Claim Processed Date Time	Pharmacy
Prescription Fill Date	Pharmacy
Issuer Claim Paid Date	Pharmacy
Prescription Service Reference Number	Pharmacy
National Drug Code	Pharmacy
Dispensing Provider ID Qualifier	Pharmacy
Dispensing Provider Identifier	Pharmacy
Prescription Fill Number	Pharmacy
Dispensing Status Code	Pharmacy
Void Replace Code	Pharmacy
Allowed Total Cost Amount	Pharmacy
Policy Paid Amount	Pharmacy
Derived Service Claim Indicator	Pharmacy
Masked Enrollee ID	Supplemental Diagnosis
Supplemental Diagnosis Detail Record Identifier	Supplemental Diagnosis
Original Claim Identifier	Supplemental Diagnosis
Detail Record Processed Date and Time	Supplemental Diagnosis
Add Delete Void Code	Supplemental Diagnosis
Original Supplemental Detail ID	Supplemental Diagnosis
Service From Date	Supplemental Diagnosis
Service To Date	Supplemental Diagnosis
Supplemental Diagnosis Code	Supplemental Diagnosis
Source Code (medical record or EDI)	Supplemental Diagnosis

4.6.3 Process and Requirements

In order to extract the previously identified variables while protecting issuer priority information and enrollee privacy, we are exploring the creation of a new EDGE server report. A new EDGE server command which would operate similarly to existing EDGE server commands (e.g. risk adjustment reports run command) would be written to create the new report. As is currently done on the EDGE server, the issuer will execute the new command when it is pushed by the EDGE Management Console. Privacy and security of the potential report data would be ensured through the use of a 256-bit encrypted string using the combination of the masked enrollee ID + Issuer ID + EDGE server ID so that the report would be encrypted once it is created and before it is sent to CMS. In this way, CMS would not know the identity of the enrollee, the issuer, or the EDGE server from which the data was extracted, allowing for the creation of a national de-identified data set.

We believe this extraction and reporting process is the least disruptive approach to accessing the most appropriate source data for recalibration while ensuring the privacy and

security of the data. The command to create the recalibration report will only require the issuer to execute the EDGE command. Furthermore, this process will require little effort on the part of the issuer since the new report will be constructed from existing EDGE server data tables and will only obligate the issuer to execute the new EDGE command without submitting new data.

4.6.4 Privacy and Security

CMS is committed to protecting the individual health information of all enrollees whose data is collected via the new EDGE report in accordance with existing privacy laws and regulations to which CMS may be subject. As covered entities under HIPAA, issuers would be required to transmit enrollee-level data as permitted by HIPAA, including the privacy and security implementing regulations to which issuers are already subject. Once the enrollee-level data is transmitted to CMS or CMS's contractors, CMS would comply with the limitations on use, storage, transport, and safeguarding of data under Federal laws and regulations, including, as applicable, the Privacy Act and the Federal Information Security Management Act (FISMA), as amended by the Federal Information Security Modernization Act.

4.6.5 Data Uses

We also consider whether this de-identified enrollee-level data could be used beyond risk adjustment recalibration. CMS would consider the following key principles in exploring other data uses:

- Improvement of CMS programs and in particular the risk adjustment program;
- Minimization of privacy and security risks; and
- Enhancement of transparency of claims and diagnosis data.

There are a number of other ways this data could be used to improve the risk adjustment program. Many concerns with the risk adjustment methodology could be addressed with the use of individual and small group market enrollee-level data for calibration and estimation. This dataset is more representative of the individual and small group market populations than commercial databases like MarketScan®, and would result in a more accurate model calibration. For example, marked differences in the incremental cost of disease X relative to disease Y between the individual and small group markets, would be accounted for if this data were used to calibrate the HHS risk adjustment models. In addition, CMS could factor in partial year enrollment, reduce the data lag on a commercial dataset and better account for high cost conditions with new treatments, and estimate factors such as CSR and induced demand more accurately. These data could also be used to assess the extent to which differences in claims costs across different types of plans can be adequately explained by the combination of the factors included in the risk adjustment model and induced demand. Those results could be used, if appropriate, to develop an additional factor for inclusion in the transfer formula that could account for higher "unobserved" risk among those enrolled in particular types of plans. CSR information could also provide a potential source for incorporating socioeconomic status into the risk adjustment model.

This dataset could also be used to improve other programs affecting individual and small group market issuers. For example, there may be opportunities for its use in the development of

the Actuarial Value (AV) Calculator, which is used to determine whether non-grandfathered individual and small group market plans on and off the Marketplaces meet actuarial value thresholds. The current claims data used in the AV Calculator has certain limitations that are unlikely to be resolved in the future, such as a lack of metal tier levels and plan network variation. We have received feedback expressing a desire for better alignment of the data underlying the AV Calculator with the individual and small group markets. A de-identified national dataset could also be used by other State or Federal governmental entities, academic institutions, or private entities to better understand private health insurance markets. However, we appreciate the need to implement any such use carefully and transparently, with full safeguards for concerns relating to privacy, security, proprietary interests. We seek comment on this approach.

CHAPTER 5: RISK ADJUSTMENT PAYMENT TRANSFER FORMULA

5.1 Introduction

As described in Chapter 1, the risk adjustment methodology has two main components: (1) a risk adjustment model or method for measuring risk selection; and (2) a method for calculating risk adjustment transfers. This chapter describes the second component. It lays out the objectives of the transfer formula, describes the transfer formula and each of its components, and presents several examples to show how risk adjustment transfers are calculated. We then describe the feedback we have received largely after the release of final risk adjustment transfers on June 30, 2015 in the "Summary Report on Transitional Reinsurance Payments and Permanent Risk Adjustment Transfers for the 2014 Benefit Year"[48] and our initial findings from the 2014 benefit year data and how various types of plans and issuers were impacted.

5.1.1 Objectives for the Transfer Formula

The purpose of risk adjustment is to facilitate the development of small group and individual markets in which consumers can select coverage based on premiums that reflect differences in plan design and benefits and not the risk of the enrollees who choose a particular plan or level of coverage. Risk adjustment transfers should reflect the risk (or health status) of each plan's enrollees, after taking into consideration its metal level, the prevailing level of expenditure in the geographic areas in which the enrollees live, the effect of coverage on utilization (induced demand), and the age and family structure of the subscriber.

We expected that risk adjustment transfers would affect the premiums plans charge. We assume issuers would seek to anticipate their transfer payments or charges. In a competitive insurance market, issuers will base premiums on revenue projections that include the financial impact of risk adjustment transfers. An issuer that anticipates experiencing adverse selection can also anticipate receiving a risk adjustment payment that will cover some or all of its higher-than-average anticipated costs; in a competitive market we expect that a risk adjustment payment will lead to lower premiums. The converse is also true; we expect that plans facing a risk adjustment charge will set higher premiums than they would in the absence of risk adjustment.

An ideal system of risk adjustment will allow premiums to reflect variation in plan costs that are the result of factors other than the health status of enrollees, including benefit design and other features of plan design such as network structure and medical management. In addition to this objective, risk adjustment transfers must meet the requirement of being budget neutral, which means that the payments and charges across an entire risk pool within a market within a State must sum to zero. These objectives and constraints guided development of the risk adjustment payment transfer formula described in the following section.

In summary, the two main objectives that the risk adjustment payment transfer formula is designed to achieve are:

---

[48] Summary Report on Transitional Reinsurance Payments and Permanent Risk Adjustment Transfers for the 2014 Benefit Year. Revised: September 17, 2015. Available at: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/RI-RA-Report-REVISED-9-17-15.pdf

1. Mitigate the effect of risk selection on plan premiums. This objective requires a calculation of how much cost variation results from actuarial risks that deviate from the market average, and imposing a payment or charge that compensates for these costs while leaving other sources of cost variation intact. Specifically, premiums should continue to reflect differences in actuarial value, geographic cost differences, age rating (within regulated limitations), and other plan characteristics.

2. Balance payments and charges across the entire Statewide market risk pool. To remain budget neutral, the sum of all charges imposed on low-risk plans should equal the payments made to high-risk plans.

In the following section, we present the risk adjustment payment transfer formula that has been designed to achieve these two objectives.

5.2 The Transfer Formula

The transfer formula is the part of the ACA risk adjustment methodology that determines the dollar flow from low- to high-risk plans. The plan liability average risk score is a crucial component of the transfer formula. But the transfer formula also incorporates several other factors into its calculation of transfers, as described below. A plan's transfer payment or charge depends on the risk scores and other information of all plans in the State individual or small group market risk pool.

5.2.1 Overview of the Transfer Formula

The purpose of risk adjustment transfers is to provide plans with enough additional revenue to cover their actual risk exposure beyond the premiums they are able to collect, or in other words, to compensate for excess actuarial risk due to risk selection. Conceptually, the transfer formula measures the difference between:

- the revenues required by a plan given the health status expenditure risk of the plan's actual enrollees; and

- the revenues that a plan can generate based on the allowable rating factors of the plan’s actual enrollees.

A plan's risk adjustment payment or charge is determined by the difference between a plan's predicted required revenues ("costs") based on its enrollees' health status and the premium revenue it can be expected to collect based on its enrollees' allowable premium rating factors.

This basic difference between required and allowable revenues is adjusted by other plan cost factors that multiplicatively interact with *both* health status expenditure risk and allowable premium revenue. Three other cost factors are modeled in this transfer payment formula:

1. the actuarial value (AV) of the plan’s benefits (metal level),

2. the plan's allowable rating factor (ARF), which reflects the relative amount a plan can charge given the age of its enrollees,

3. a geographic cost factor (GCF).

With these adiustments. the formula is:

With these adjustments, the formula is:

$$T_i = \left[\frac{PLRS_i \cdot IDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)} - \frac{AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)}\right] \bar{P}_s$$

The first term in the transfer formula within the square brackets defines the revenue required by a plan (relative to the Statewide market average). It is the product of three terms:

1. a plan liability risk score (PLRS), which reflects the plan's actuarial value as well as the plan's enrollee health status risk,[49]

2. an induced demand factor (IDF), which reflects the anticipated induced demand associated with the plan's cost sharing (metal) level, and

3. a geographic cost factor (GCF), which reflects prevailing utilization and expenditure patterns in the geographic location of the plan's enrollees.

The second term in the transfer formula within the square brackets defines the revenue that a plan can be expected to generate given the allowable rating factors (relative to the Statewide market average). It is the product of four terms:

1. the actuarial value (AV) of the plan's benefits (metal level),

2. the plan's allowable rating factor (ARF), which reflects the relative amount a plan can charge given the age of its enrollees,

3. the induced demand factor (IDF) associated with the plan's metal level, and

4. the geographic cost factor (GCF) of the plan's enrollees.

The denominators of both the first and second terms in the transfer equation express Statewide average required revenue and allowable premium, respectively. Dividing the first and second terms by the respective Statewide average expresses the plan's revenue requirement and allowable revenue relative to the average plan offered in the State. Transfers are converted from relative factors into dollar amounts by multiplying them by the Statewide enrollment-weighted market average plan premium $\bar{P}_S$.

The transfer formula assumes a multiplicative relationship among the various cost factors. Other things being equal, a 10 percent increase in the cost of doing business in a rating area increases plan liabilities and premiums by 10 percent, a 10 percent increase in risk increases plan liabilities by 10 percent, etc. If Plan A's actuarial value is 25 percent higher than Plan B's AV, and Plan A's geographic cost factor is 40 percent higher than Plan B's GCF, then Plan A's costs would be expected to be 75 percent greater than Plan B (1.25*1.40 – 1.00 = 1.75).

[49] The risk score is also multiplicatively adjusted for induced demand associated with individual enrollee income-based Affordable Care Act cost sharing subsidies.

More formally, a plan's expected premium requirement (cost) is assumed to be proportional to the product of the Plan Liability Risk Score, the Induced Demand Factor, and the Geographic Cost Factor:

$$\left[\frac{PLRS_i \cdot IDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)}\right] \bar{P}_S$$

where

$\bar{P}_S$ = Statewide market average premium $[\bar{P}_S = \sum_i (s_i \cdot P_i)]$

P_i = Average premium per member month of plan i

$PLRS_i$ = plan i's plan liability risk score,

IDF_i = plan i's induced demand factor,

GCF_i = plan i's geographic cost factor,

s_i = plan i's share of marketwide enrollment,

and the denominator is summed across all plans in the Statewide market risk pool.

Conceptually, this expression calculates the plan's expected cost or revenue requirement relative to the marketwide average costs. The denominator of the adjustment term normalizes the product of the plan cost factors to the market average product of these cost factors. This normalized product of the plan cost factors is an estimate of how underlying differences in a plan's cost factors – specifically risk selection, induced demand and geographic cost differences – cause the plan's liability to differ from that of the average plan offered in the market. So for example, if the term in brackets is equal to 1.12, this indicates that based solely on these three factors, the plan's cost would be 12 percent higher than those of the average plan. This measure of expected costs is then scaled by the Statewide market average premium $(\bar{P}_S)$ to estimate the premium revenue that would be necessary to cover the plan's expected costs.

The key factor in this premium estimate is the plan liability risk score, which is calculated from the HHS risk adjustment model. The PLRS is a relative measure of plan liability based on the health status of the plan's enrollees. This risk score measure incorporates each plan's actuarial value (based on metal level), and implicitly reflects the higher proportion of enrollees' costs that will be covered by plans with a higher actuarial value. Because of this, the required premium expression does not include a separate AV adjustment factor.

The other expression in the transfer formula simulates how much premium revenue the plan would be expected to collect. This expression is shown below:

$$\left[\frac{AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)}\right] \bar{P}_S$$

where the notation is as defined above, and

ARF_i = the allowable rating factor of plan i.

The ARF reflects the impact of the age composition of each plan's enrollees on the premium it would collect from subscribers, given applicable rating constraints.

This term is an estimate of the amount of premium revenue that a plan can be expected to collect from its subscribers with premiums that are based solely on age. The AV, IDF and GCF adjusters are needed because these factors affect the plan's operating costs independently of enrollee age or risk. Plans with higher actuarial value, induced demand, or geographic area costs would be expected to charge higher premiums to cover these costs. The ARF adjustment captures the fact that plans will charge older enrollees higher premiums, though the age-based premium is restricted by applicable rating rules. Note that the right hand allowable revenue term does not include a plan's risk score, because health status risk is not an allowable rating factor under the ACA. This term reflects the relationship between the premium that the plan could collect from enrollees if they were of average risk but taking into account their age, geographic location, the actuarial value of their coverage, and the induced demand associated with that level of coverage.

The Statewide market average premium acts as a common scaling factor for both terms in the formula, both of which are expressed relative to the Statewide market average. The Statewide average premium will also reflect the Statewide cost level. Over the long run, the Statewide average premium is expected to equal the Statewide average cost (including allowable loading for administrative costs, surplus, and profit). The Statewide premium is therefore simultaneously a premium and a cost scaling factor. The Statewide average premium embeds an average level of efficiency. All plans receive a risk adjustment payment or charge sufficient for a plan with average efficiency.

Two other reasons that transfers are scaled by the Statewide average premium, as opposed to, for example, the plan's own premium, are:

- Using the Statewide average premium minimizes issuers' ability to manipulate their transfers by adjusting their own plan premiums.

- Scaling all transfers to the same premium, combined with the assumption that the factors affecting premium requirements and allowable revenue have a multiplicative relationship, obviates any further adjustment of payments and charges to ensure that risk adjustment transfers for the entire market sum to zero.[50]

Structured as shown above, the transfer formula calculates T_i, the payment or (if negative) charge to plan i for each member month of enrollment. The total transfer for each plan is calculated by multiplying T_i by the plan's total billable member months.

For the purposes of the transfer formula, a health plan that is offered in more than one geographic rating area is treated as if it was comprised of separate plans, one offered exclusively in each rating area. The risk score, geographic cost factor, and other elements of the transfer formula interact multiplicatively. For this reason, multi-rating-area plans must be treated as

[50] There would still need to be a pro rata adjustment to payments should charges not be paid in full by some issuers. However, mathematically, payments will by definition equal charges using the methods described in this section.

separate "plan segments," one per rating area, to calculate transfers correctly. Once transfers have been calculated for each plan segment at the rating area level, they can be re-aggregated to the plan or issuer level.

5.2.2 Why the Formula Uses a Multiplicative Model for Plan Costs

As discussed above, the two terms in the transfer formula – the required premium (PLRS*IDF*GCF) and the realized premium (AV*ARF*IDF*GCF) – both reflect a multiplicative relationship among the various cost factors. The transfer formula assumes that each of the various cost factors are independently and proportionately related to revenue requirements and allowable premium. That is, other things being equal (risk, induced demand), a 10 percent increase in the prevailing level of costs in a rating area will increase both revenue requirements (plan costs) and allowable premium revenue by 10 percent. Similarly, within a geographic area, a 10 percent increase in induced demand will increase both revenue requirements and allowable premium revenue by 10 percent.

To see the effect of this assumed multiplicative relationship, suppose a plan's enrollees carry health risks equal to the market wide average, and these costs are $500 per member per month. If it is a gold plan, these costs are expected to be 8 percent higher, or $540. Furthermore, if the plan operates in a geographic area where costs are 50 percent higher than average, we would expect these additional costs to affect not only the base $500, but rather the entire $540, bringing the plan's expected costs to $810 (=$540*150 percent). In other words, the high costs in this geographic area apply not just to the baseline level of costs, but also to the cost variations from other sources, in this case the induced utilization linked to being a gold plan.

5.3 Specific Adjustments: Concepts, Issues, and Measurement

In this section we provide details on the various components of the transfer formula.[51] For each variable, we describe its role in the transfer formula, explain how the variable is measured, and address any issues that arise relating to it.

5.3.1 Risk Scores: Plan Liability versus Total Expenditure

The plan liability risk average score (PLRS) represents the plan's overall risk exposure. It incorporates differences in actuarial value as reflected in each plan's assigned metal level. Each individual has a different PLRS depending on what metal tier of plan he enrolls in.[52] The relationship between plan actuarial value and plan liability is non-linear, because of the presence of deductibles and out-of-pocket limits. Hence, the PLRS cannot be obtained by simply multiplying a person's total expenditure risk score by his or her plan's actuarial value. In addition, the PLRS includes an adjustment to account for the family rating rules described in the Market Reform Rule (78 Federal Register 13405), which caps (at three) the number of children who will count toward the buildup of family rates.

---

[51] More information on each of the components of the transfer formula, and on the calculation of transfers, is available in the 2014 Payment Notice (78 Federal Register 15410).

[52] In contrast, a person's total expenditure risk score is the same regardless of plan, because expected total expenditures do not vary based on a plan's cost sharing rules.

The formula below shows how the PLRS is calculated, including the adjustment for family rating rules:

$$PLRS_i = \frac{\sum_e M_e \cdot PLRS_e}{\sum_b M_b}$$

where

$PLRS_i$ is plan i's average plan liability risk score,

the subscript e denotes each enrollee within the plan,

$PLRS_e$ is each enrollee's individual plan i liability risk score,

M_e is the number of months during the risk adjustment period the enrollee e is enrolled in the plan i, and

M_b is the number of months during the risk adjustment period the billable enrollee b is enrolled in plan i (as described in more detail below, billable enrollees exclude children who do not count towards the premium).

Each individual enrollee's PLRS is calculated from the applicable risk adjustment model. These metal level-specific risk adjustment models produce risk scores that reflect enrollee health status as well as the AV of the plan.

The numerator of the PLRS formula calculates a sum that reflects the risk score of every person covered by the plan. The Federal rules for family rating allow an issuer to charge a premium only for up to three children. A family with four or more children would continue to accrue claims that reflect the expected cost of all family members. The additional children would not, however, contribute to the premium revenue of the issuer. Including only all billable enrollees in the denominator, (excluding children who do not count toward family policy premiums) causes the weighted average PLRS to take into account the fact that families with non-billed children impose more risk per billed member month than families in which every member month is billable, all else being equal.

Other than this formula for the PLRS, all other calculations in the transfer formula are performed using the number of billable member months of enrollment.

5.3.2 Actuarial Value

The actuarial value (AV) adjustment in the transfer formula accounts for relative differences in plan liability due to differences in the percentage of enrollees' expenditures that the plan covers.[53] It is, however, explicitly included only in the allowable premium revenue. Actuarial value is, as noted above, implicitly reflected in the PLRS.

Without an AV adjustment, low-risk, low-AV plans would tend to pay higher charges than appropriate (because their claims expense would not be scaled down by their low AV), and

[53] Recall that the AV adjustment is implicitly made in the plan liability risk score in the first term of the transfer formula, and is explicitly made as a separate component in the right hand side of the transfer formula.

high-risk, high-AV plans would receive lower payments than are necessary to compensate for their excess risk (because their payments would not be scaled up by their high AV). Concomitantly, high-risk, low-AV plans would tend to receive higher payments than appropriate and low-risk, high AV plans would pay less charges than appropriate.

Importantly, the AV reflects the percentage of total expenditures that the plan will cover if it enrolls a "standard" or "normal" population, not the percentage of the actual enrollee expenditures that the plan will cover. Actuarial value interacts with risk such that the percentage of expenditures that the plan covers will likely be lower for a low-risk than for a high-risk population.

The PLRS implicitly reflects the actuarial value of the plan. The PLRS empirical model predicts simulated plan liability using diagnoses and demographics. Plan liability is simulated using benefit parameters on a standard population (the entire MarketScan® sample) that yield the metal level AV. The AV implicit in the PLRS is thus consistent with the explicit AV in the second term of the transfer formula.

5.3.3 Induced Demand

Induced demand (or induced utilization) reflects differences in enrollee spending patterns attributable to differences in the generosity of plan benefits (cost sharing). Risk adjustment should not compensate issuers for plan liability attributed to variation in benefit design. For this reason, the transfer formula includes an induced demand adjustment to the plan revenue requirement and allowable premium revenue sides of the equation. The induced demand factors for each metal level are used for the values of IDF_i in the payment transfer formula.[54]

5.3.4 Allowable Rating Factors

The Affordable Care Act specifies four enrollee characteristics that issuers are allowed to use as rating factors: enrollee age, family size, tobacco use, and geographic area within the State. The statute further limits the degree to which premiums can vary by age; among adults (age 21 and up), the most expensive age group's rating cannot be more than three times as high as the lowest. Basing premiums on enrollee health status is explicitly prohibited.

The allowable rating factor (ARF) adjustment in the transfer formula accounts only for age rating. The risk score is adjusted for family rating requirements as described above, and geographic rating areas receive a separate adjustment also described below. Tobacco use and wellness discounts are not included in the ARF.[55]

---

[54] The induced demand factor in the transfer formula was calculated via an analysis of the expected difference in expenditures for expected enrollees in risk adjustment covered plans of different actuarial values. For this analysis, the factors were derived through HHS analysis and appear to be broadly consistent with results from the RAND Health Insurance Experiment (Newhouse, 1984).

[55] Tobacco rating and wellness discounts are discretionary, and are not included in the ARF to maintain issuer flexibility regarding these rating adjustments. For example, including an adjustment in the transfer formula for tobacco use could create an incentive for issuers to rate based on tobacco usage. A tobacco use adjustment would have the effect of lowering the transfer payments received for tobacco users in a pool of enrollees, which could prompt issuers to rate for tobacco.

Age rating allows issuers to be compensated for risk selection through premiums paid by subscribers to the extent that it is based on and reflects enrollee age. Under Federal rules, each State has a standard age curve that all plans are required to use (a Federal age rating curve operates in States that do not designate their own curve, referred to as the Affordable Care Act default age curve). The 3:1 age rating restriction applies to adults only, defined as age 21 or greater. Each plan's ARF is calculated as the enrollment-weighted average of age factors across all of the plan's enrollees.

For example, a 21-year-old enrollee has an ARF of 1.000, while the maximum rating (for people age 64 and older) is 3.000, conforming to the 3:1 maximum adult age rating restriction. Plan-level average ARF values enable comparisons of premium-generating capacity across plans; for example, a plan whose ARF is 2.00 would be able to collect 25 percent more premium revenue through age rating than a plan with an ARF equal to 1.60. The ARF can be modified to conform to State variation in rating rules (e.g., a 2:1 rather than 3:1 age rating restriction with associated age rating curve).

Under the Market Reform Rule, premiums for families with three or fewer children are calculated as the sum of individual rates for each individual within the family. These individual rates are based on each person's age, tobacco use, and geographic rating area. For families with more than three children, the family premium would be built from individual premiums of the parents plus the three oldest children. Additional children would not be reflected in the family premium.

We note that some States use family tiering rating factors (New York and Vermont). In family tiering States, family tiering rating factors are not required to yield premiums that are equal to the sum of the individual policy members' applicable rates, nor must they be set in a way that counts only the rates of the oldest three children under 21 within a family policy. To account for the differences in family rating practices between family tiering States and non-family tiering States, we finalized a modification to the ARF formula for family tiering States in the second Program Integrity Rule (78 Federal Register 65056) and a further clarification in the 2016 Payment Notice (80 Federal Register 10750). This modification modified the ARF formula for use in these States so that the numerator is a summation over all subscribers of the product of the family tiering factor and the subscriber member months, and the denominator the sum of billable member months.

As discussed earlier, however, average plan liability risk scores do take family size into account by including the actuarial risks of non-billable family members in the calculation of the average over all billable enrollees.

5.3.5 Geographic Cost Variation

The transfer formula includes an adjustment for geographic cost variation because there are many costs, such as input prices and medical care utilization rates that vary geographically and are likely to affect premiums. Without a geographic cost adjustment, risk adjustment transfers would tend to distort the transfer payment or charge. These distortions would include:

- In low cost areas, high-risk plans would, other things being equal, receive a larger transfer payment than is needed to mitigate the effect of unfavorable selection.

- In a high cost area, high-risk plans would receive a transfer payment that is insufficient to mitigate the effect of unfavorable selection.
- Low-risk plans in high-cost areas would, other things being equal, be assessed a smaller transfer charge than is required to mitigate the effect of favorable selection.
- A low-risk plan in a high-cost area would, by the same token, be assessed a larger charge than is needed to mitigate the effect of favorable selection.

A geographic cost factor (GCF) is calculated for each rating area. The purpose of the GCF is to reflect differences in utilization patterns and the cost of doing business across geographic areas, *but not to reflect differences in risk across areas*. To achieve this goal, for the metal level risk pool,[56] the GCF is calculated based on the observed average silver plan premiums in a geographic area relative to the Statewide average silver plan premium. All issuers must offer a silver plan in any rating area in which they offer any plan. Using only silver plans as the basis of the adjustment controls for differences in average premium revenue (i.e., the amount paid by subscribers for coverage) across rating areas that are attributable to differences in actuarial value and induced demand.

The average premium payment for a rating area will reflect the area's geographic rating factor, but will not reflect differences in risk between rating areas. As a result, the silver plan premiums used to calculate the adjustment are standardized for age rating to isolate geographic cost differences embedded in premiums. They do not need to be standardized for risk score because observed premiums, which already include the anticipated payment transfers, are implicitly adjusted for differences in risk.

Calculation of the GCF involves three steps. First, each issuer calculates the average premium for each silver plan it offers in each rating area. As discussed above, for purposes of these calculations a single plan that is offered in multiple rating areas is broken into segments and treated as consisting of separate plans in each rating area in which it is offered.[57] The calculation for a single rating area is:

$$\overline{P}_i = \frac{\sum_s (M_s \cdot P_s)}{\sum_b M_b}$$

where

$\overline{P}_i$ is the average premium for plan i;

s indexes all subscribers enrolled in the plan;

M_s is the number of billable member months for the subscriber s;

P_s is the premium for subscriber s; and

[56] For the catastrophic risk pool, the GCF is calculated off of catastrophic premiums.

[57] Plans offered in multiple areas must be decomposed into plan segments to support the calculation of the geographic cost factors and to recognize the multiplicative relationship among the factors included in the transfer formula.

M_b is the number of billable members b enrolled in the plan.

The second step is to standardize the average premium revenue for each plan offered in a rating area for the effect of age. Plan premiums are standardized for age by dividing the average plan premium by the plan's allowable rating factor. This formula is:

$$\overline{P}_i^{AS} = \overline{P}_i/(ARF_i)$$

where

$\overline{P}_i^{AS}$ is plan i's age standardized average premium;

$\overline{P}_i$ is the average premium for plan i; and

ARF_i is plan i's allowable rating factor.

The third step is to compute a GCF for each area. The GCF is simply the ratio of the enrollment-weighted average age-standardized premium revenue for a rating area to the overall Statewide enrollment-weighted average age-standardized premium revenue for all silver plans. This calculation of the GCF for a rating area is:

$$GCF_i = \left(\sum_{area} s_i^a \cdot \overline{P}_i^{AS}\right) \Big/ \left(\sum_{state} s_i^s \cdot \overline{P}_i^{AS}\right)$$

where

s_i^a is the share of plan i's enrollment in area a, and

s_i^s is the share of plan i's area a enrollment as a portion of all enrollments in the State.

This equation divides the enrollment-weighted average of standardized silver plan premiums in a geographic area by the average of those premiums Statewide. The numerator's summation is over all silver plans within plan i's geographic area, so $\sum_{area} s_i^a = 1$. Similarly, the summation in the denominator is over all silver plans in the State, so $\sum_{state} s_i^s = 1$.

Using these formulas, the enrollment-weighted Statewide average of plan GCF values will equal 1.0, so the GCF can be interpreted as the percentage by which any geographic area's costs deviate from the State average. In other words, a GCF equal to 1.15 indicates that the plan operates in a geographic area where costs are, on average, 15 percent higher than the Statewide average.

Splitting/Aggregating Plans Across Geographic Areas

If a plan enrolls members in multiple rating areas, it will be decomposed into "plan segments" with enrollment (and any other characteristics necessary for the transfer calculation that vary by area) specific to each applicable rating area. These plan segments will be used in the calculation of payment transfers, which will then be aggregated to the plan and issuer level for execution.

During the development of the transfer formula, a question arose whether it is necessary to calculate risk adjustment transfers separately for each plan-geographic area combination, or if it is possible to calculate transfers at the level of the plan. Disaggregation of premium data to the level of the rating area is required because the transfer amount is determined by the product of the factors that are included in the transfer formula. Calculating transfers using the product of the enrollment-weighted average factors (i.e., the enrollment-weighted average risk score, the enrollment-weighted average induced demand factor, etc.) will yield incorrect transfers because the product of the averages does not equal the average of the products. Risk adjustment will be performed by calculating transfers separately in each geographic area in which a plan operates.

5.3.6 Family Size and Member Months of Enrollment (the "Billable" Concept)

The transfer formula is based on individuals (including children) rather than families as the unit of enrollment. However, as described earlier, the PLRS includes an adjustment to account for the family rating rules, which cap (at 3) the number of children who can count toward the buildup of family rates. When the plan average PLRS is calculated, all plan enrollees are counted in the numerator, but only billable plan enrollees (parents and up to three oldest children) are counted in the denominator. This creates a weighted average plan PLRS that takes into account the fact that families with non-billable children impose more risk per billable member month than families in which every member month is billable, all else being equal.

Thus, the added actuarial risks generated by large families (those with more than three dependent children) are taken into account by the average plan liability risk score. At the same time, plans can cover these additional costs by building them into the premiums charged per billable member month. The transfer formula calculates a transfer payment or charge per billable member month, and both terms in the formula are consistent with this unit of measurement. Risks and premiums are measured per billable member month.

In addition, the transfer formula can accommodate State variation in rating rules. As long as the PLRS is calculated as total enrolled risks divided by total billable member months, and the plan's average ARF is calculated from all billable individuals' ARF values, the transfer formula will compensate for cost variations due to enrollee health status risks per billable member month. This value (T_i), multiplied by the number of billable member months of enrollment, will approximate the costs of excess actuarial risks for a typical plan in the State, with adjustments for differences in geographic area costs, actuarial value, and induced demand.

5.4 Other Analytical Issues/Features

5.4.1 Normalization of Multiplied Factors

One feature of the transfer formula is that it can accommodate any scaling or normalization of individual adjustment factors. Any of the relative measures that enter the formula – a list that includes PLRS, AV, IDF, GCF, and ARF – can be rescaled without affecting the resulting transfers.

This invariance to the scaling or normalization of individual adjustment factors allows flexibility in how these factors are measured. Risk scores in particular are usually normalized to a 1.0 mean, and it is helpful that the transfer formula yields the same results regardless of whether this normalization is imposed.

The normalization is essentially built in to the transfer formula in the denominators of the two terms in brackets. Each term (PLRS*IDF*GCF and AV*ARF*IDF*GCF) is divided by the market-wide weighted average of this product, which allows each term to equal plan i's total adjustment relative to the State average adjustment.

Note that transfers *cannot* accommodate the rescaling of two factors in the transfer formula: the plan market share s_i and the Statewide average premium $\overline{P}_s$. The plan market shares appear only in the denominators of the transfer formula terms in brackets and must sum to 1.0 so that the denominators represent the Statewide average of the product of the plan cost factors. The Statewide average premium provides the dollar scale to the transfers; the transfers will vary proportionally with the Statewide average premium.

5.4.2 How Risk Adjustment Transfers are Budget Neutral (Balanced)

By construction, risk adjustment payments and charges will be budget neutral, meaning the total amount of risk adjustment charges collected from issuers will equal the total amount of risk adjustment payments made. This is a key objective of the transfer formula, and it ensures that risk adjustment will neither require government sources of funding nor accumulate a surplus.

Risk adjustment transfers are budget neutral because the enrollment-weighted means of both terms in the formula are equal to each other. The total transfer across all providers is given by:

$$T = \sum_i \left\{ s_i \cdot \overline{P}_s \cdot \left[\frac{PLRS_i \cdot IDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)} - \frac{AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)} \right] \right\}$$

Because $\overline{P}_s$ is a constant for all plans i, and both of the terms inside the square brackets are multiplied by the same factor, s_i, this can be restated:

$$T = \overline{P}_s \cdot \sum_i \left\{ s_i \cdot \left[\frac{PLRS_i \cdot IDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)} \right] - s_i \cdot \left[\frac{AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)} \right] \right\}$$

Because the sum of the difference between two variables is equal to the difference between the sums of the two variables:

$$T = \overline{P}_s \cdot \left\{ \sum_i \left[\frac{s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)} \right] - \sum_i \left[\frac{s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)} \right] \right\}$$

Because $\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)$ and $\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)$ are constants for all plans i, this can be further restated:

$$T = \overline{P}_s \cdot \left\{ \left[\frac{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)} \right] - \left[\frac{\sum_i (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)} \right] \right\}$$

Which can be simplified:

$$T = \overline{P}_s \cdot \{1.0 - 1.0\}$$

$$T = \overline{P}_s \cdot 0$$

$$T = 0$$

Note that each of the two premium adjustment terms in the formula is normalized (divided) by the weighted average of these adjustments across the entire State. By definition, this normalization ensures that the weighted average of each adjustment is exactly 1, and the enrollment-weighted sum will be the same for both adjustments. As a result, one weighted sum subtracted from the other will equal zero.

5.5 Evaluation of 2014 Benefit Year Risk Adjustment Results and Additional Proposals

Following the release of the June 30, 2015 "Summary Report on Transitional Reinsurance Payments and Permanent Risk Adjustment Transfers for the 2014 Benefit Year,"[58] whereby issuers were notified of 2014 benefit year risk adjustment transfer amounts, we received feedback from issuers that some HHS risk adjustment transfers negatively affected some issuers. Some of the feedback suggested that CMS should reconsider including administrative costs that are not related to risk in the Statewide average premium, for example. We received comments that risk adjustment transfers were unfair to small issuers, new market entrants, and fast-growing issuers, with recommendations to exempt these issuers, phase-in these issuers, and/or to cap risk adjustment transfers for all issuers.

5.5.1 Including Administrative Costs in the Statewide Average Premium

As discussed above, the Statewide average premium is intended to reflect average administrative expenses and average claims costs for issuers in a market and State. We received comments from the public who believe that the inclusion of administrative costs in the Statewide average premium incorrectly increases risk adjustment transfers based on costs that are unrelated to the risk of the enrollee population. Comments ranged from requesting that administrative expenses be removed entirely from the Statewide average premium to requesting that HHS consider basing risk adjustment transfers on a portion of Statewide average premium – namely, the portion representing the sum of claims, claims adjustment expenses, and taxes that are calculated on premiums after risk adjustment transfers by using a specified percentage of Statewide average premiums. One suggestion noted that a specified percentage could be determined based on data submitted by issuers on the Unified Rate Review Template (URRT) for the portion of premium needed for claims and on data from financial reporting statements for claims adjustment expenses and relevant taxes as a percent of premium.

We note that low cost plans do not necessarily indicate efficient plans. Should a plan be low cost with low claims costs, it is likely an indication of mispricing, as the issuer should be pricing for average risk. However, we understand the concern that including fixed administrative costs in the Statewide average premium may increase risk adjustment transfers for all issuers based on a percentage of costs that are not related to enrollee risk. We considered some of the potential effects of excluding fixed administrative costs that are unrelated to enrollee risk from the Statewide average premium. This modification to the treatment of administrative costs in the Statewide average premium would lower absolute risk adjustment transfers for all

[58] Summary Report on Transitional Reinsurance Payments and Permanent Risk Adjustment Transfers for the 2014 Benefit Year. Revised: September 17, 2015. Available at: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/RI-RA-Report-REVISED-9-17-15.pdf

issuers by an equal percentage. We also note that administrative costs are affected by claims costs and that correctly measuring the portion of administrative costs unaffected by claims costs may be difficult. An incorrect measurement of administrative costs could then result in plans with high risk enrollees being undercompensated. We are continuing to evaluate the impact of administrative expenses on risk adjustment transfers, and may consider this adjustment beyond the 2018 benefit year.

5.5.2 Potential Change to the Transfer Formula

We are also investigating whether the risk adjustment methodology appropriately addresses plan differences not fully captured by aspects of the current risk adjustment methodology. For example, although a number of sources of premium variation – such as metal level, age, and geographic cost factors – are explicitly addressed in the transfer equation, others – such as network differences, plan efficiency, or effective care coordination or disease management – are not. We are exploring a number of ways of addressing such plan differences in our methodology, including through potentially modifying the transfer equation, perhaps by modifying the equation using a plan's own premium, though we are cognizant of potential sources of error from inaccurate rate setting as well as the risks of creating unintended incentives to raise or lower premiums in order to take advantage of this effect. We welcome comments and analysis on whether and how we might make such a change to the methodology to address some of these effects.

5.5.3 Risk Adjustment Issuers for 2014 Benefit Year

CMS implemented a distributed data collection process consisting of an External Data Gathering Environment (EDGE) located at each of 775 issuers, connected to a central management server. EDGE server deployment commenced September 2013 and data collection started upon deployment, depending on issuer readiness to submit to their server. CMS monitored data completeness on each server – namely, we used enrollment and claims baseline estimates from issuers used to evaluate data completion. To determine data submission success, we used a benchmark of 90 percent of enrollment and claims data as compared to an issuer's baseline data.[59] We also monitored key data quality indicators using an empirical outlier analysis:

- Issuers with outliers either explained the issue as a unique population characteristic or as a data extract/submission problem; and
- Problematic data had to be corrected or the issuer would receive a default risk adjustment charge and/or forego reinsurance payments.

CMS ensured that the data used within risk adjustment risk pools to calculate risk adjustment transfers met these thresholds, but even still, some issuers maintain that they had data processing errors and issues that prevented their plan liability risk score from being higher, and thus resulting in a lower risk adjustment charge, or a higher risk adjustment payment.

[59] Evaluation of EDGE Data Submissions Bulletin for the 2014 Benefit Year. April 24, 2015. Available at: https://www.cms.gov/cciio/resources/regulations-and-guidance/

We implemented these data quality and quantity checks consistently with all issuers. Some issuers who have stated they were overcharged had more difficulty than others in achieving quality data submissions, which could have affected risk adjustment transfer amounts, though CMS considers these difficulties to be preventable by the issuer itself. We also believe that some issuers were unsure of how to rate their plans to account for risk adjustment in the initial year of the program, which also would have affected risk adjustment payments and charges. It is likely that pricing will become more accurate with each year of the program.

This section describes the outcomes of risk adjustment in the individual and small group markets in 2014. The distribution of risk transfers as a percentage of plan and issuer premium revenue is analyzed. Then, risk transfers by plan metal level are described, followed by an examination of risk transfers by issuer size.

The analysis is conducted at a national level, and includes all plans subject to risk adjustment. Massachusetts, which conducts its own risk adjustment, is excluded.[60] The District of Columbia is included as a "State." The market subject to risk adjustment includes both plans offered through the State-based Marketplaces and Federally-facilitated Marketplaces and plans offered off-Marketplace. Unless otherwise indicated, all analyses in this section were conducted with data from the EDGE Calculation Module, using the June 24, 2015 version. The source of the data is the Risk Adjustment Transfer Elements Extract (RATEE) report from each issuer's EDGE server. The EDGE data are used by CMS in risk adjustment operations to calculate risk transfers. EDGE data are available at the State, market segment, issuer, plan, and rating area level.

The analyses in this section are aggregated to the State, market segment, issuer and plan, or State, market segment and issuer level. When reading and interpreting the data contained in the tables, the following concepts and terminology should be kept in mind:

- An issuer is the corporate entity that offers health plans in a market. An issuer may offer a single plan or may offer multiple plans (differing in terms of benefit design or other plan features) in each "metal level."
- A plan is defined by its benefit design and other features such as provider network. A plan may be offered in a single "rating area" or may be offered in multiple "rating areas." The premium for the plan will vary by rating area.
- A rating area is the geographic area within which the plan is offered. The premium charged for the plan will vary by rating area.

5.5.4 Risk Adjustment Transfer Analysis for 2014 Benefit Year

We examined how the size of risk adjustment transfers varied across issuers in the individual market. Of the 468 issuers in the risk adjustment eligible individual market, 200

[60] We did not recertify the alternate risk adjustment methodology in Massachusetts for the 2017 benefit year, and as such, HHS will operate risk adjustment in all States in 2017.

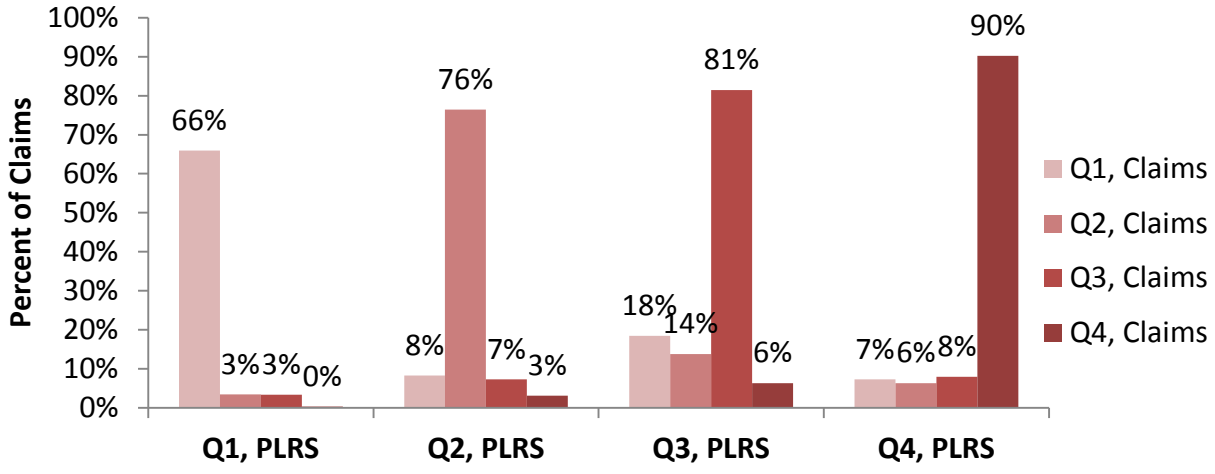
issuers (43 percent) owed charges, while 268 issuers (57 percent) received payments.[61] Our analysis has identified three key themes from the 2014 benefit year experience:

- Both the HHS risk adjustment model and risk adjustment transfers were highly related to an issuer's relative amount of paid claims;
- An issuer's size did not predetermine if they received a risk adjustment payment or charge; and
- Differences in issuers' mix of enrollees across metal levels were related to the direction and magnitude of risk adjustment transfers, but was not the only determining factor.

Risk Adjustment and Paid Claims

Consistent with the rationale of the risk adjustment program issuers with relatively low paid claims per enrollee mostly had relatively low plan liability risk scores and issuers with relatively high paid claims per enrollee tended to have relatively high risk scores (**Figure 5.1**). That is, issuers with relatively low paid claims per enrollee (issuers in the first quartile when ranked by per enrollee claims) owed risk adjustment charges of 17 percent of premium, on average. On the other hand, issuers with relatively large paid claims per enrollee (those ranked in the fourth quartile by per enrollee claims) received payments that constituted 9 percent of premium, on average.[62]

Figure 5.1: Distribution of issuers by quartile of PLRS and quartile of per enrollee claims, Individual Market



Amounts shown are at the issuer level and weighted by billable MMs of enrollment.

[61] Data tabulated by Agency for Healthcare Research and Quality (AHRQ) as of March 1, 2016.

[62] Data tabulated by AHRQ as of March 1, 2016.

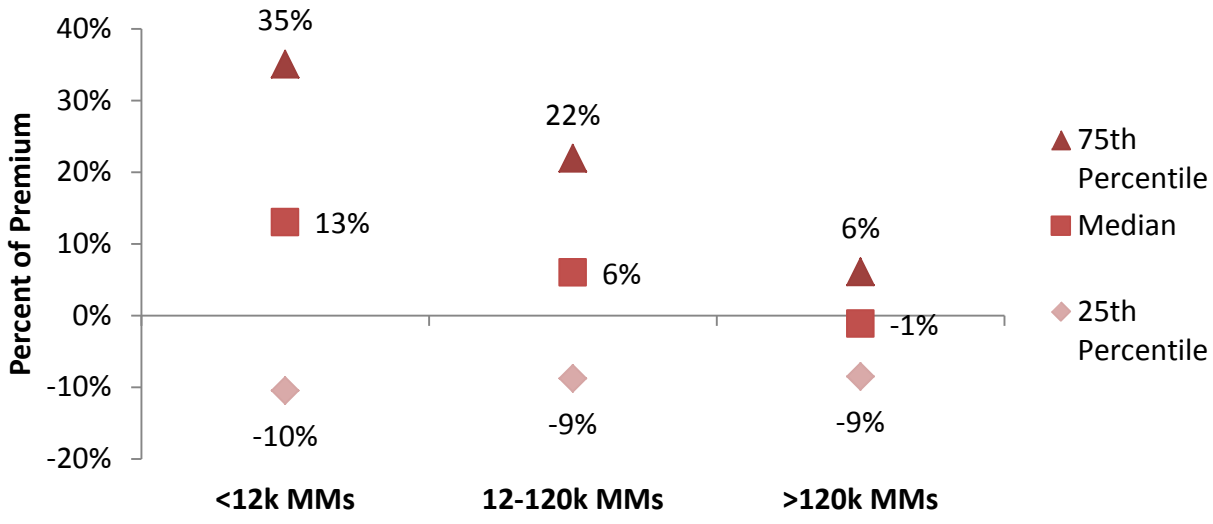
Issuers in the lowest quartile of paid claims per enrollee owed the most risk adjustment charges as a percentage of their premiums, and issuers in the top quartile received the most risk adjustment payments as a share of premiums. Issuers with lower paid claims per enrollee tended to pay in more to risk adjustment, and issuers in the top quarter of paid claims received risk adjustment payments as a share of premiums, with a median of 2 percent of premium. For issuers with relatively low paid claims per enrollee (issuers in the first quartile when ranked by per enrollee claims), the 25th percentile of transfers to the risk adjustment program was a charge of 21 percent of premiums and a charge of 6 percent at the 75th percentile, on average. For issuers with relatively large paid claims per enrollee (those ranked in the fourth quartile by per enrollee claims), the 25th percentile of transfers had a charge of 1 percent of their premiums, while the 75th percentile received a payment of 10 percent. Accounting for reinsurance payments and risk adjustment transfers, among issuers ranked in the first quartile of paid claims per enrollee (low paid claims per enrollee), the median issuer owed 15 percent of their premiums in risk adjustment charges, and after accounting for reinsurance payments, the median issuer owed net of 2 percent of premium. The median issuer in the highest quartile of paid claims per enrollee received risk adjustment and reinsurance payments comprising 29 percent of premiums. Additionally, the amount of combined risk adjustment payments and charges, and reinsurance payments as a percentage of issuer premium increases moving from the lowest to the highest quartile of issuer paid claims amounts.[63]

Risk Adjustment and Issuer Size

In the individual market, on average, smaller issuers received risk adjustment payments while larger issuers owed risk adjustment charges. However, there was substantial variability in payments and charges particularly among smaller issuers. Risk adjustment transfers as a share of issuer premiums varied much less for larger issuers (those with more than 120,000 billable member months) than for smaller issuers (those with less than 12,000 billable member months). Median transfers in the individual market as a share of premiums were net payments for smaller issuers (13percent and 6 percent) and a small charge amount (-1 percent) in the largest issuer category (those with greater than 120,000 billable member months) (**Figure 5.2**). Small group risk adjustment transfers as a share of issuer premiums varied much less for larger issuers (those with more than 120,000 billable member months) than for smaller issuers (those with less than 12,000 billable member months) (**Figure 5.3**). Median transfers as a share of premiums were net charges for smaller issuers and a small net payment (1 percent) for the largest issuers (those with greater than 120,000 billable member months).

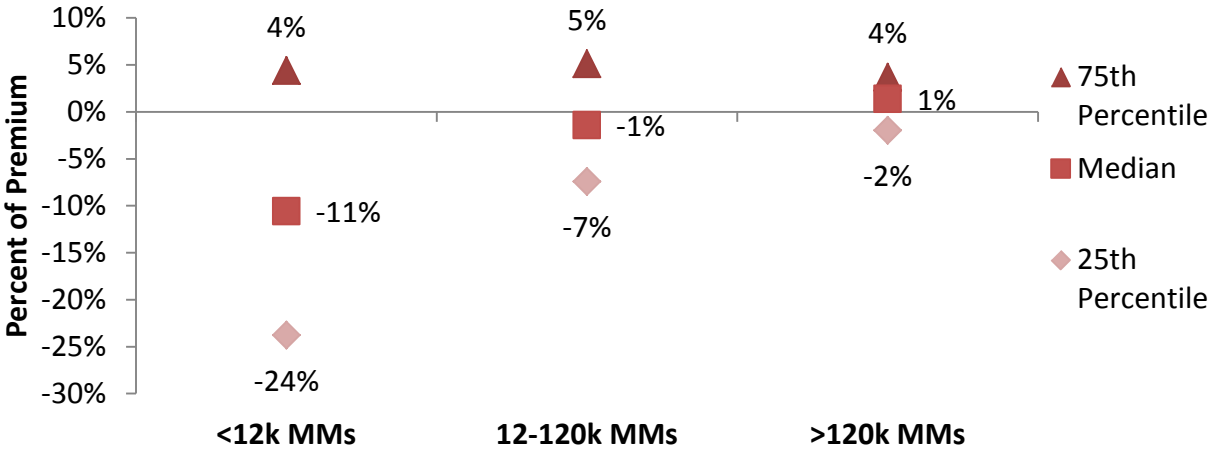
[63] Data tabulated by AHRQ as of March 1, 2016.

Figure 5.2: Distribution of transfers as percent of premium, by issuer size, Individual Market



Amounts shown at issuer level and weighted by billable members months of enrollment.

Figure 5.3: Distribution of transfers as percent of premium, by issuer size, Small Group Market



Amounts shown at issuer level and weighted by billable members months of enrollment.

Risk Adjustment and Metal Level Distribution

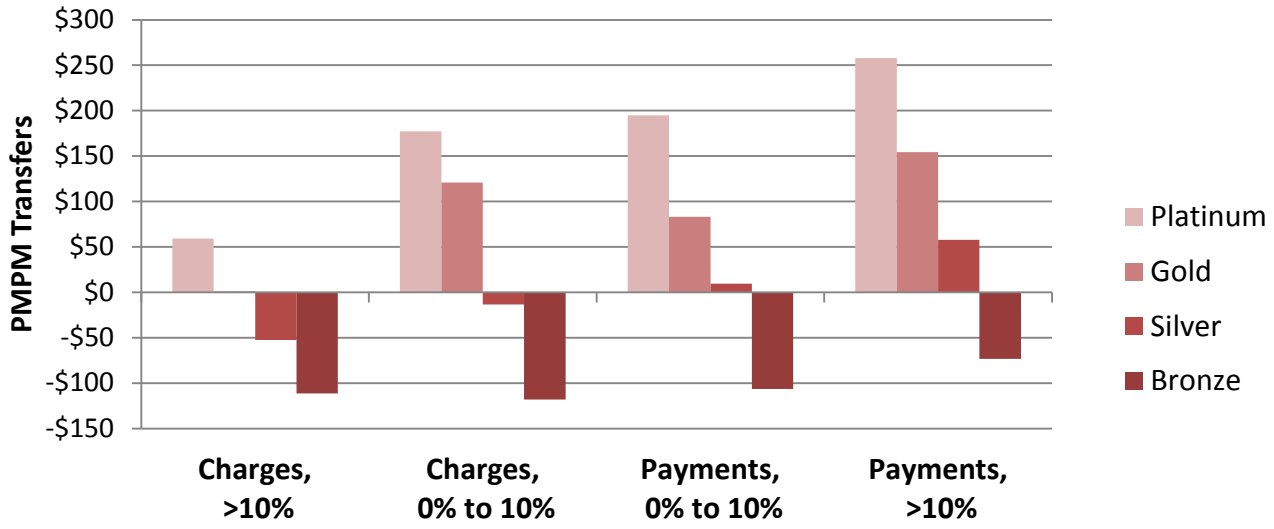
The distribution of enrollees across metal levels is related to the issuers' amount of payments or charges – issuers with greater than average proportions of enrollees in gold and platinum plans are more likely to receive payments, and issuers with greater than average proportions of members in bronze plans are more likely to owe charges. Issuers in the individual market that owed large risk adjustment charges as a share of premiums (those owing total charges greater than 10 percent of premium) tended to have large charges for their bronze plans and relatively small payments for their platinum or gold plans (**Figure 5.4**). Issuers in the other three categories of risk adjustment transfers as a share of premiums generally received relatively large risk adjustment payments for their platinum and gold plans. In particular, issuers that received risk adjustment payments in excess of 10 percent of their premiums tended to receive very large risk adjustment payments for their platinum and gold plans and relatively large payments for silver plans as well. Small group issuers that owed large risk adjustment charges as a share of premiums (those owing total charges greater than 10 percent of premium) tended to have large charges across all metal levels (**Figure 5.5**). Issuers with other risk adjustment transfers as a share of premiums were generally owed relatively large amounts for their platinum plans.

Differences in issuers' mix of enrollees across metal levels were related to the direction and magnitude of risk adjustment transfers, but they were not the only determining factor. Treating each individual market issuer equally (not weighting by enrollment), enrollment shares across metal levels explained about one-quarter (27 percent) of the variation in issuer-level risk scores. Treating each small group issuer equally (not weighting by enrollment), enrollment shares across metal levels only explained about one-quarter (22 percent) of the variation in issuer-level risk scores.[64, 65]

[64] Data tabulated by Agency for Healthcare Research and Quality (AHRQ) as of March 1, 2016.

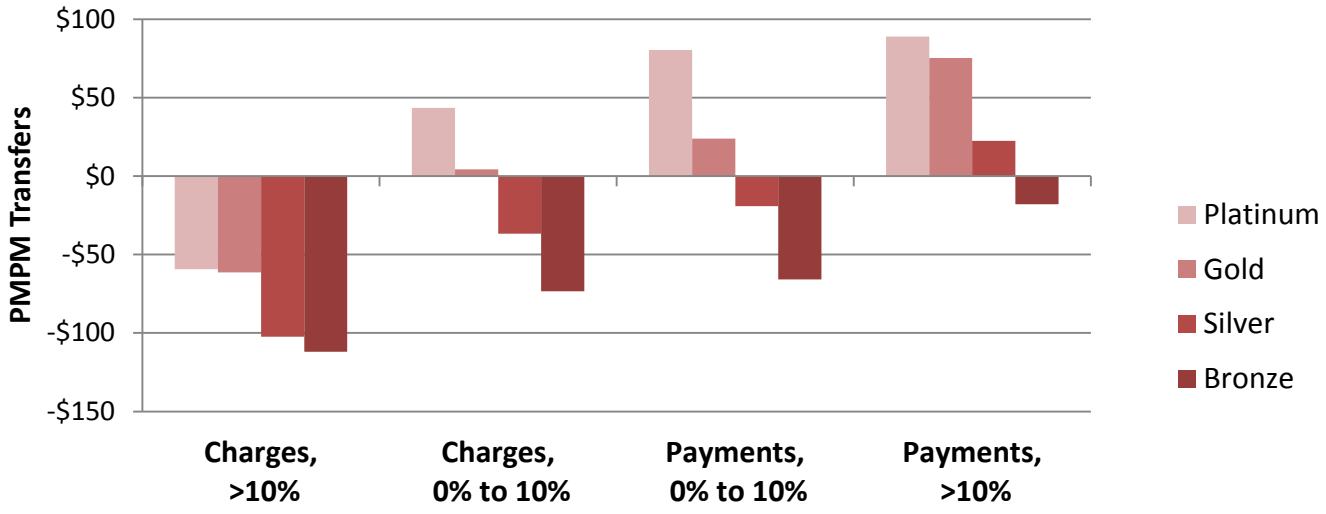
[65] To assess the degree to which issuers' scores varied because of enrollment in different metal level tiers, we performed regressions of issuer risk scores relative to the State market average on the shares of each issuer's enrollment in the various metal levels. The variation in the predictions from this regression was about 22 percent and 41 percent as large as the actual variability in relative risk scores in the unweighted and weighted regressions, respectively.

Figure 5.4: RA PMPM transfers by metal level and by transfers as percent of issuer revenues, Individual Market



All calculations at the issuer level and weighted by member months.

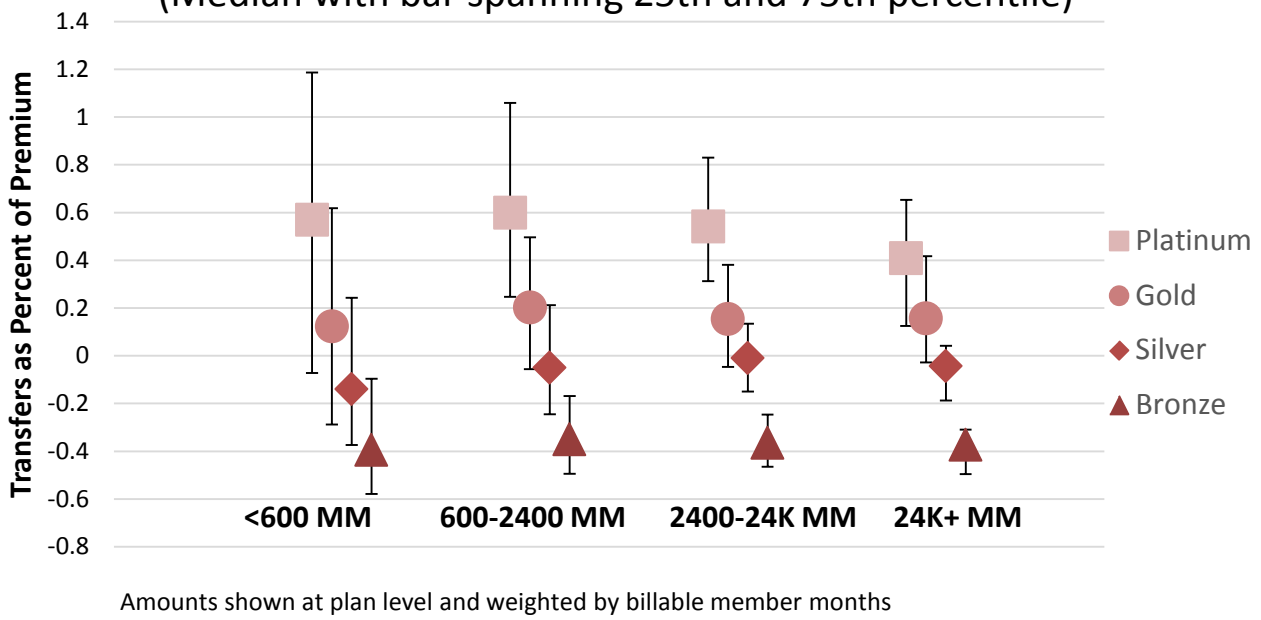
Figure 5.5: RA PMPM transfers by metal level and by transfers as percent of issuer revenues, Small Group Market



All calculations at the issuer level and weighted by member months.

Individual market plans with smaller enrollment experienced noticeably larger variation in risk adjustment transfers as a percentage of premiums. However, bronze plans, regardless of enrollment size, consistently owed risk adjustment charges while platinum plans across all enrollment sizes generally received payments (**Figures 5.8**). Across all metal levels, median small group plans with smaller enrollment owed risk adjustment charges and experienced noticeably larger variation in risk adjustment transfers as a percentage of premiums (**Figure 5.9**).[66]

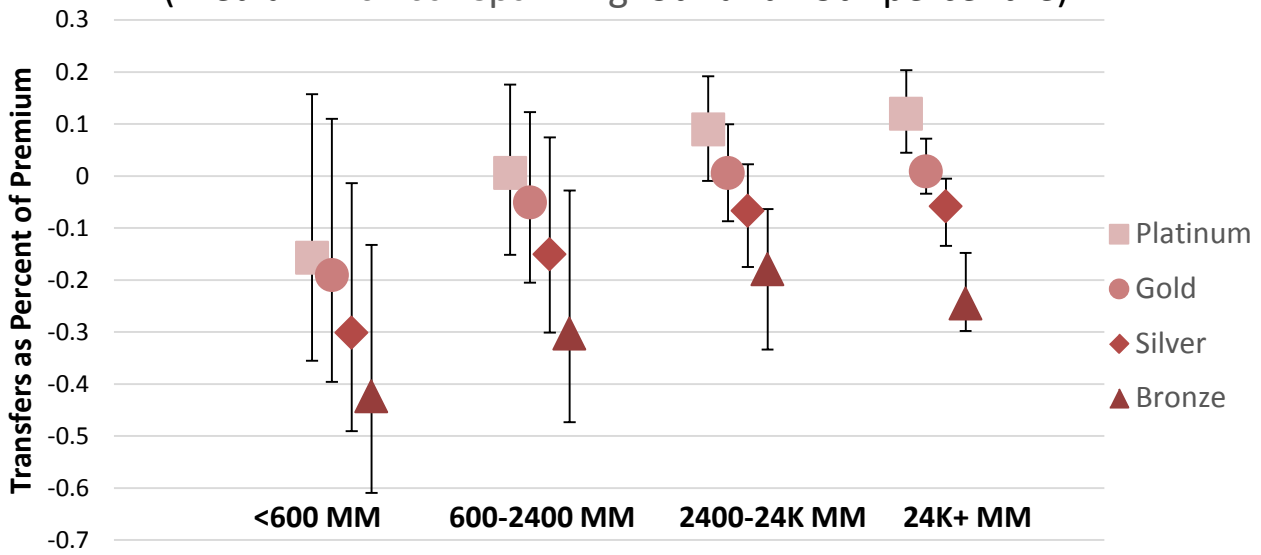
Figure 5.8: Distribution of transfers as percent of premium, by metal level and plan size, Individual Market
(Median with bar spanning 25th and 75th percentile)



Amounts shown at plan level and weighted by billable member months

[66] Data tabulated by AHRQ as of March 1, 2016.

Figure 5.9: Distribution of transfers as percent of premium, by metal level and plan size, Small Group Market
(Median with bar spanning25th and 75th percentile)



Amounts shown at plan level and weighted by billable member months

5.6 Summary and Conclusions

CMS has implemented a risk adjustment program to mitigate the effects of risk selection on health insurance premiums for non-grandfathered plans in the individual and small group markets. The risk adjustment program, supports market stability by pooling risk and transferring funds from plans with more low-risk (i.e., healthier and lower cost) enrollees to those plans with more high-risk (i.e., less healthy and higher cost) enrollees. The initial findings from benefit year 2014 indicate that, in general, the HHS risk adjustment methodology is working as intended. We look forward to feedback from the public on the considerations presented in this paper and anticipate this feedback informing valuable improvements to the HHS risk adjustment methodology for the 2018 benefit year and beyond.

REFERENCES

American Academy of Actuaries. Risk Adjustment and Risk Assessment. Issue Brief, May 2010.

Armstrong, J., Paolucci, F., McLeod, H., van de Ven, W.: Risk Equalization in Voluntary Health Insurance Markets: A Three Country Comparison. Health Policy 98 (2010): 39-49, 2011.

Barry, C.L., Weiner, J.P., Lemke, K., et al.: Risk Adjustment in Health Insurance Exchanges for Individuals with Mental Illness. Am J Psych 169(7):704-709, July 2012.

Buettgens, M., Garrett, B., and Holahan, J.: American Under the Affordable Care Act. Urban Institute and Robert Wood Johnson Foundation, December 2010.

Centers for Disease Control and Prevention (CDC): International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). Available at http://www.cdc.gov/nchs/icd/icd9cm.htm

Centers for Disease Control and Prevention (CDC): International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). Available at http://www.cdc.gov/nchs/icd/icd10cm.htm

Department of Health and Human Services: Patient Protection and Affordable Care Act; Actuarial Value Calculator Methodology. Available at http://www.cms.gov/CCIIO/Resources/Files/Downloads/av-calculator-methodology.pdf

Kautter, J., Ingber, M., Pope, G. C., & Freeman, S.: Improvements in Medicare Part D risk adjustment: Beneficiary access and payment accuracy. Medical Care 50:1102–1108, 2012.

Medicare Payment Advisory Commission (MedPAC), 2012a. Medicare Payment Basics: Medicare Advantage Program Payment System, 2012.

Medicare Payment Advisory Commission (MedPAC), 2012b. Medicare Payment Basics: Medicare Part D Payment System, 2012.

Newhouse, J.P., Brook, R. H., Ware, J. E., et al.: The Effect of Coinsurance on the Health of Adults: Results from the RAND Health Insurance Experiment. Santa Monica, Calif.: RAND Corporation, R-3055-HHS, December 1984.

Patient Protection and Affordable Care Act, 2010. Pub. L. No. 111-148, March 23, 2010.

Pope, G.C., Kautter, J., Ellis, R.P., et al.: Risk Adjustment for Medicare Capitation Payments Using the CMS-HCC Model. Health Care Financing Review 25(4):119-141, Summer, 2004.

Schokkaert, E., Beck, K., Shmueli, A., van de Ven, W., van de Voorde, C., and Wasem, J.: Acceptable Costs and Risk Adjustment: Policy Choices and Ethical Trade-Offs. Katholieke Universiteit Leuven, Department of Economics, Discussion Paper, November 2006.

Trish, E., Damico, A., Claxton, G. et al.: A Profile of Health Insurance Exchange Enrollees. The Kaiser Family Foundation, March 2011.

Winkelman, R., and Damler, R.: Risk Adjustment in State Medicaid Programs. Health Watch. Society of Actuaries, January 2008.

APPENDIX

Appendix A: 2017 Risk Adjustment Model Factors

Table A.1:
Adult Risk Adjustment Model Factors

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Demographic Factors					
Age 21-24, Male	0.236	0.180	0.119	0.082	0.081
Age 25-29, Male	0.246	0.186	0.122	0.083	0.082
Age 30-34, Male	0.287	0.216	0.138	0.089	0.088
Age 35-39, Male	0.346	0.264	0.172	0.112	0.111
Age 40-44, Male	0.420	0.326	0.221	0.151	0.149
Age 45-49, Male	0.496	0.392	0.273	0.192	0.191
Age 50-54, Male	0.633	0.512	0.372	0.275	0.274
Age 55-59, Male	0.722	0.585	0.429	0.320	0.318
Age 60-64, Male	0.843	0.683	0.502	0.372	0.369
Age 21-24, Female	0.379	0.296	0.200	0.138	0.137
Age 25-29, Female	0.460	0.359	0.247	0.173	0.172
Age 30-34, Female	0.582	0.466	0.337	0.254	0.252
Age 35-39, Female	0.668	0.542	0.405	0.318	0.316
Age 40-44, Female	0.742	0.604	0.455	0.357	0.355
Age 45-49, Female	0.750	0.608	0.450	0.344	0.342
Age 50-54, Female	0.845	0.691	0.518	0.398	0.395
Age 55-59, Female	0.849	0.690	0.510	0.380	0.378
Age 60-64, Female	0.909	0.734	0.537	0.395	0.392
. .		*Diagnosis Factors*			
HIV/AIDS	8.942	8.450	8.099	8.142	8.143
Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	10.686	10.511	10.405	10.461	10.462
Central Nervous System Infections, Except Viral Meningitis	6.632	6.532	6.468	6.489	6.489
Viral or Unspecified Meningitis	4.657	4.422	4.263	4.222	4.222
Opportunistic Infections	8.503	8.404	8.337	8.319	8.319
Metastatic Cancer	24.314	23.880	23.578	23.637	23.638

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia	12.630	12.296	12.062	12.066	12.066
Non-Hodgkin`s Lymphomas and Other Cancers and Tumors	5.845	5.611	5.435	5.388	5.387
Colorectal, Breast (Age < 50), Kidney, and Other Cancers	5.152	4.918	4.738	4.690	4.689
Breast (Age 50+) and Prostate Cancer, Benign/Uncertain Brain Tumors, and Other Cancers and Tumors	2.957	2.786	2.650	2.597	2.596
Thyroid Cancer, Melanoma, Neurofibromatosis, and Other Cancers and Tumors	1.448	1.295	1.160	1.069	1.067
Pancreas Transplant Status/Complications	5.455	5.233	5.091	5.112	5.114
Diabetes with Acute Complications	1.187	1.049	0.925	0.822	0.820
Diabetes with Chronic Complications	1.187	1.049	0.925	0.822	0.820
Diabetes without Complication	1.187	1.049	0.925	0.822	0.820
Protein-Calorie Malnutrition	13.686	13.693	13.702	13.762	13.763
Mucopolysaccharidosis	2.277	2.159	2.061	2.008	2.007
Lipidoses and Glycogenosis	2.277	2.159	2.061	2.008	2.007
Amyloidosis, Porphyria, and Other Metabolic Disorders	2.277	2.159	2.061	2.008	2.007
Adrenal, Pituitary, and Other Significant Endocrine Disorders	2.277	2.159	2.061	2.008	2.007

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Liver Transplant Status/Complications	16.042	15.868	15.759	15.771	15.772
End-Stage Liver Disease	7.119	6.877	6.718	6.736	6.737
Cirrhosis of Liver	3.852	3.690	3.569	3.535	3.535
Chronic Hepatitis	3.852	3.690	3.569	3.535	3.535
Acute Liver Failure/Disease, Including Neonatal Hepatitis	4.430	4.269	4.158	4.148	4.148
Intestine Transplant Status/Complications	32.604	32.555	32.516	32.559	32.559
Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis	11.820	11.561	11.383	11.413	11.413
Intestinal Obstruction	6.537	6.272	6.101	6.120	6.121
Chronic Pancreatitis	5.455	5.233	5.091	5.112	5.114
Acute Pancreatitis/Other Pancreatic Disorders and Intestinal Malabsorption	2.702	2.515	2.379	2.331	2.331
Inflammatory Bowel Disease	3.657	3.392	3.190	3.098	3.096
Necrotizing Fasciitis	6.576	6.378	6.239	6.254	6.255
Bone/Joint/Muscle Infections/Necrosis	6.576	6.378	6.239	6.254	6.255
Rheumatoid Arthritis and Specified Autoimmune Disorders	4.848	4.587	4.394	4.385	4.385
Systemic Lupus Erythematosus and Other Autoimmune Disorders	1.205	1.070	0.952	0.868	0.867
Osteogenesis Imperfecta and Other Osteodystrophies	3.115	2.917	2.758	2.699	2.697
Congenital/Developmental Skeletal and Connective Tissue Disorders	3.115	2.917	2.758	2.699	2.697
Cleft Lip/Cleft Palate	1.295	1.137	1.010	0.942	0.941

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Hemophilia	46.436	46.150	45.931	45.939	45.939
Myelodysplastic Syndromes and Myelofibrosis	12.671	12.534	12.440	12.448	12.449
Aplastic Anemia	12.671	12.534	12.440	12.448	12.449
Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn	9.737	9.576	9.454	9.445	9.445
Sickle Cell Anemia (Hb-SS)	9.737	9.576	9.454	9.445	9.445
Thalassemia Major	9.737	9.576	9.454	9.445	9.445
Combined and Other Severe Immunodeficiencies	5.432	5.284	5.182	5.183	5.183
Disorders of the Immune Mechanism	5.432	5.284	5.182	5.183	5.183
Coagulation Defects and Other Specified Hematological Disorders	2.805	2.707	2.628	2.599	2.599
Drug Psychosis	3.830	3.574	3.380	3.286	3.284
Drug Dependence	3.830	3.574	3.380	3.286	3.284
Schizophrenia	3.189	2.934	2.744	2.680	2.679
Major Depressive and Bipolar Disorders	1.714	1.547	1.404	1.308	1.307
Reactive and Unspecified Psychosis, Delusional Disorders	1.714	1.547	1.404	1.308	1.307
Personality Disorders	1.176	1.043	0.910	0.814	0.812
Anorexia/Bulimia Nervosa	2.693	2.527	2.392	2.334	2.333
Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes	2.632	2.504	2.403	2.354	2.353
Down Syndrome, Fragile X, Other Chromosomal Anomalies, and Congenital Malformation Syndromes	1.056	0.951	0.849	0.778	0.776

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Autistic Disorder	1.176	1.043	0.910	0.814	0.812
Pervasive Developmental Disorders, Except Autistic Disorder	1.176	1.043	0.910	0.814	0.812
Traumatic Complete Lesion Cervical Spinal Cord	12.005	11.851	11.737	11.735	11.735
Quadriplegia	12.005	11.851	11.737	11.735	11.735
Traumatic Complete Lesion Dorsal Spinal Cord	9.157	9.000	8.886	8.874	8.875
Paraplegia	9.157	9.000	8.886	8.874	8.875
Spinal Cord Disorders/Injuries	5.635	5.424	5.275	5.246	5.246
Amyotrophic Lateral Sclerosis and Other Anterior Horn Cell Disease	3.029	2.792	2.625	2.585	2.585
Quadriplegic Cerebral Palsy	1.206	0.997	0.839	0.777	0.776
Cerebral Palsy, Except Quadriplegic	0.124	0.068	0.034	0.011	0.011
Spina Bifida and Other Brain/Spinal/Nervous System Congenital Anomalies	0.071	0.019	0.000	0.000	0.000
Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy	5.247	5.099	4.994	4.971	4.971
Muscular Dystrophy	2.147	1.981	1.860	1.785	1.784
Multiple Sclerosis	13.590	13.187	12.905	12.950	12.951
Parkinson`s, Huntington`s, and Spinocerebellar Disease, and Other Neurodegenerative Disorders	2.147	1.981	1.860	1.785	1.784

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Seizure Disorders and Convulsions	1.495	1.337	1.207	1.137	1.136
Hydrocephalus	6.388	6.266	6.165	6.139	6.138
Non-Traumatic Coma, and Brain Compression/Anoxic Damage	9.207	9.070	8.964	8.958	8.957
Respirator Dependence/Tracheostomy Status	34.719	34.708	34.706	34.772	34.773
Respiratory Arrest	10.554	10.403	10.306	10.370	10.371
Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes	10.554	10.403	10.306	10.370	10.371
Heart Assistive Device/Artificial Heart	35.114	34.869	34.711	34.771	34.772
Heart Transplant	35.114	34.869	34.711	34.771	34.772
Congestive Heart Failure	3.280	3.171	3.095	3.089	3.089
Acute Myocardial Infarction	10.129	9.795	9.580	9.691	9.693
Unstable Angina and Other Acute Ischemic Heart Disease	5.227	4.952	4.779	4.793	4.794
Heart Infection/Inflammation, Except Rheumatic	6.297	6.163	6.063	6.042	6.041
Specified Heart Arrhythmias	2.829	2.681	2.565	2.512	2.511
Intracranial Hemorrhage	9.423	9.144	8.954	8.963	8.964
Ischemic or Unspecified Stroke	3.167	2.982	2.869	2.875	2.876
Cerebral Aneurysm and Arteriovenous Malformation	3.940	3.742	3.600	3.559	3.558
Hemiplegia/Hemiparesis	5.468	5.374	5.317	5.360	5.361
Monoplegia, Other Paralytic Syndromes	3.452	3.319	3.226	3.207	3.207

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Atherosclerosis of the Extremities with Ulceration or Gangrene	10.940	10.840	10.784	10.853	10.854
Vascular Disease with Complications	7.727	7.543	7.416	7.417	7.417
Pulmonary Embolism and Deep Vein Thrombosis	3.841	3.675	3.555	3.529	3.529
Lung Transplant Status/Complications	36.419	36.227	36.103	36.180	36.181
Cystic Fibrosis	18.011	17.687	17.444	17.467	17.467
Chronic Obstructive Pulmonary Disease, Including Bronchiectasis	0.942	0.825	0.717	0.641	0.640
Asthma	0.942	0.825	0.717	0.641	0.640
Fibrosis of Lung and Other Lung Disorders	1.889	1.771	1.682	1.641	1.640
Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections	7.594	7.520	7.471	7.485	7.485
Kidney Transplant Status	10.183	9.919	9.744	9.735	9.735
End Stage Renal Disease	38.463	38.228	38.078	38.198	38.201
Chronic Kidney Disease, Stage 5	2.088	1.989	1.925	1.920	1.920
Chronic Kidney Disease, Severe (Stage 4)	2.088	1.989	1.925	1.920	1.920
Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism	1.340	1.156	0.979	0.795	0.791
Miscarriage with Complications	1.340	1.156	0.979	0.795	0.791
Miscarriage with No or Minor Complications	1.340	1.156	0.979	0.795	0.791

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Completed Pregnancy With Major Complications	3.630	3.150	2.862	2.712	2.713
Completed Pregnancy With Complications	3.630	3.150	2.862	2.712	2.713
Completed Pregnancy with No or Minor Complications	3.630	3.150	2.862	2.712	2.713
Chronic Ulcer of Skin, Except Pressure	2.356	2.233	2.150	2.134	2.134
Hip Fractures and Pathological Vertebral or Humerus Fractures	9.460	9.245	9.100	9.136	9.136
Pathological Fractures, Except of Vertebrae, Hip, or Humerus	2.000	1.871	1.758	1.688	1.687
Stem Cell, Including Bone Marrow, Transplant Status/Complications	31.027	31.022	31.017	31.035	31.036
Artificial Openings for Feeding or Elimination	10.038	9.946	9.886	9.924	9.925
Amputation Status, Lower Limb/Amputation Complications	5.263	5.112	5.015	5.044	5.045
Interaction Factors					
Severe illness x Opportunistic Infections	10.408	10.632	10.799	10.894	10.895
Severe illness x Metastatic Cancer	10.408	10.632	10.799	10.894	10.895
Severe illness x Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia	10.408	10.632	10.799	10.894	10.895

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Severe illness x Non-Hodgkin`s Lymphomas and Other Cancers and Tumors	10.408	10.632	10.799	10.894	10.895
Severe illness x Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy	10.408	10.632	10.799	10.894	10.895
Severe illness x Heart Infection/Inflammation, Except Rheumatic	10.408	10.632	10.799	10.894	10.895
Severe illness x Intracranial Hemorrhage	10.408	10.632	10.799	10.894	10.895
Severe illness x HCC group G06 (G06 is HCC Group 6 which includes the following HCCs in the blood disease category: 67, 68)	10.408	10.632	10.799	10.894	10.895
Severe illness x HCC group G08 (G08 is HCC Group 8 which includes the following HCCs in the blood disease category: 73, 74)	10.408	10.632	10.799	10.894	10.895
Severe illness x End-Stage Liver Disease	1.906	2.039	2.141	2.225	2.226
Severe illness x Acute Liver Failure/Disease, Including Neonatal Hepatitis	1.906	2.039	2.141	2.225	2.226

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Severe illness x Atherosclerosis of the Extremities with Ulceration or Gangrene	1.906	2.039	2.141	2.225	2.226
Severe illness x Vascular Disease with Complications	1.906	2.039	2.141	2.225	2.226
Severe illness x Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections	1.906	2.039	2.141	2.225	2.226
Severe illness x Artificial Openings for Feeding or Elimination	1.906	2.039	2.141	2.225	2.226
Severe illness x HCC group G03 (G03 is HCC Group 3 which includes the following HCCs in the musculoskeletal disease category: 54, 55)	1.906	2.039	2.141	2.225	2.226

Table A.2:
Child Risk Adjustment Model Factors

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Demographic Factors					
Age 2-4, Male	0.224	0.145	0.067	0.021	0.020
Age 5-9, Male	0.155	0.098	0.038	0.004	0.004
Age 10-14, Male	0.220	0.158	0.089	0.053	0.053
Age 15-20, Male	0.290	0.219	0.142	0.097	0.096
Age 2-4, Female	0.178	0.109	0.044	0.011	0.010
Age 5-9, Female	0.127	0.076	0.027	0.003	0.002
Age 10-14, Female	0.204	0.145	0.085	0.054	0.054
Age 15-20, Female	0.330	0.248	0.157	0.101	0.100
Diagnosis Factors					

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
HIV/AIDS	4.875	4.437	4.110	4.033	4.032
Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	17.228	17.069	16.969	16.994	16.995
Central Nervous System Infections, Except Viral Meningitis	10.808	10.631	10.506	10.511	10.511
Viral or Unspecified Meningitis	3.128	2.925	2.775	2.687	2.686
Opportunistic Infections	22.943	22.880	22.834	22.825	22.825
Metastatic Cancer	36.648	36.404	36.207	36.207	36.207
Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia	12.117	11.833	11.604	11.547	11.546
Non-Hodgkin`s Lymphomas and Other Cancers and Tumors	9.328	9.058	8.836	8.754	8.753
Colorectal, Breast (Age < 50), Kidney, and Other Cancers	3.508	3.291	3.097	2.989	2.987
Breast (Age 50+) and Prostate Cancer, Benign/Uncertain Brain Tumors, and Other Cancers and Tumors	3.016	2.816	2.642	2.538	2.537
Thyroid Cancer, Melanoma, Neurofibromatosis, and Other Cancers and Tumors	1.723	1.553	1.397	1.294	1.292
Pancreas Transplant Status/Complications	30.468	30.333	30.245	30.256	30.256
Diabetes with Acute Complications	2.521	2.197	1.946	1.703	1.699
Diabetes with Chronic Complications	2.521	2.197	1.946	1.703	1.699
Diabetes without Complication	2.521	2.197	1.946	1.703	1.699
Protein-Calorie Malnutrition	13.570	13.484	13.421	13.450	13.450
Mucopolysaccharidosis	8.509	8.238	8.020	7.987	7.986
Lipidoses and Glycogenosis	8.509	8.238	8.020	7.987	7.986

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Congenital Metabolic Disorders, Not Elsewhere Classified	8.509	8.238	8.020	7.987	7.986
Amyloidosis, Porphyria, and Other Metabolic Disorders	8.509	8.238	8.020	7.987	7.986
Adrenal, Pituitary, and Other Significant Endocrine Disorders	8.509	8.238	8.020	7.987	7.986
Liver Transplant Status/Complications	30.468	30.333	30.245	30.256	30.256
End-Stage Liver Disease	13.077	12.927	12.822	12.821	12.821
Cirrhosis of Liver	9.604	9.445	9.326	9.286	9.286
Chronic Hepatitis	2.567	2.418	2.280	2.216	2.215
Acute Liver Failure/Disease, Including Neonatal Hepatitis	12.729	12.576	12.460	12.447	12.447
Intestine Transplant Status/Complications	30.468	30.333	30.245	30.256	30.256
Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis	14.795	14.463	14.217	14.238	14.238
Intestinal Obstruction	5.389	5.155	4.965	4.885	4.884
Chronic Pancreatitis	9.713	9.478	9.319	9.319	9.319
Acute Pancreatitis/Other Pancreatic Disorders and Intestinal Malabsorption	2.561	2.426	2.303	2.217	2.216
Inflammatory Bowel Disease	6.321	5.943	5.650	5.553	5.551
Necrotizing Fasciitis	4.467	4.231	4.041	3.989	3.988
Bone/Joint/Muscle Infections/Necrosis	4.467	4.231	4.041	3.989	3.988
Rheumatoid Arthritis and Specified Autoimmune Disorders	3.904	3.662	3.448	3.365	3.364
Systemic Lupus Erythematosus and Other Autoimmune Disorders	1.305	1.154	1.003	0.893	0.891
Osteogenesis Imperfecta and Other Osteodystrophies	1.560	1.429	1.303	1.232	1.231
Congenital/Developmental Skeletal and Connective Tissue Disorders	1.560	1.429	1.303	1.232	1.231

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Cleft Lip/Cleft Palate	1.563	1.351	1.172	1.061	1.059
Hemophilia	66.792	66.309	65.939	65.927	65.927
Myelodysplastic Syndromes and Myelofibrosis	15.978	15.807	15.672	15.654	15.654
Aplastic Anemia	15.978	15.807	15.672	15.654	15.654
Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn	7.706	7.432	7.214	7.145	7.144
Sickle Cell Anemia (Hb-SS)	7.706	7.432	7.214	7.145	7.144
Thalassemia Major	7.706	7.432	7.214	7.145	7.144
Combined and Other Severe Immunodeficiencies	6.686	6.507	6.364	6.310	6.309
Disorders of the Immune Mechanism	6.686	6.507	6.364	6.310	6.309
Coagulation Defects and Other Specified Hematological Disorders	4.828	4.689	4.560	4.494	4.493
Drug Psychosis	5.390	5.135	4.948	4.887	4.887
Drug Dependence	5.390	5.135	4.948	4.887	4.887
Schizophrenia	5.242	4.853	4.561	4.472	4.471
Major Depressive and Bipolar Disorders	1.913	1.691	1.485	1.334	1.332
Reactive and Unspecified Psychosis, Delusional Disorders	1.913	1.691	1.485	1.334	1.332
Personality Disorders	0.783	0.653	0.504	0.376	0.374
Anorexia/Bulimia Nervosa	2.742	2.539	2.370	2.309	2.308
Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes	3.362	3.155	3.013	2.980	2.979
Down Syndrome, Fragile X, Other Chromosomal Anomalies, and Congenital Malformation Syndromes	1.787	1.605	1.459	1.378	1.376
Autistic Disorder	1.771	1.577	1.389	1.248	1.246
Pervasive Developmental Disorders, Except Autistic Disorder	0.907	0.766	0.597	0.448	0.445

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Traumatic Complete Lesion Cervical Spinal Cord	13.209	13.168	13.154	13.225	13.227
Quadriplegia	13.209	13.168	13.154	13.225	13.227
Traumatic Complete Lesion Dorsal Spinal Cord	11.619	11.410	11.267	11.269	11.270
Paraplegia	11.619	11.410	11.267	11.269	11.270
Spinal Cord Disorders/Injuries	4.847	4.614	4.433	4.359	4.358
Amyotrophic Lateral Sclerosis and Other Anterior Horn Cell Disease	8.218	7.979	7.791	7.744	7.744
Quadriplegic Cerebral Palsy	3.387	3.141	2.983	2.995	2.996
Cerebral Palsy, Except Quadriplegic	0.861	0.675	0.530	0.451	0.450
Spina Bifida and Other Brain/Spinal/Nervous System Congenital Anomalies	1.282	1.135	1.010	0.944	0.943
Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy	9.635	9.457	9.315	9.279	9.279
Muscular Dystrophy	3.374	3.176	3.021	2.948	2.947
Multiple Sclerosis	8.431	8.101	7.852	7.820	7.820
Parkinson`s, Huntington`s, and Spinocerebellar Disease, and Other Neurodegenerative Disorders	3.374	3.176	3.021	2.948	2.947
Seizure Disorders and Convulsions	2.095	1.913	1.735	1.609	1.607
Hydrocephalus	5.122	5.002	4.912	4.903	4.903
Non-Traumatic Coma, and Brain Compression/Anoxic Damage	7.539	7.391	7.276	7.236	7.235
Respirator Dependence/Tracheostomy Status	40.112	40.012	39.969	40.084	40.086

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Respiratory Arrest	12.354	12.151	12.015	12.013	12.013
Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes	12.354	12.151	12.015	12.013	12.013
Heart Assistive Device/Artificial Heart	30.468	30.333	30.245	30.256	30.256
Heart Transplant	30.468	30.333	30.245	30.256	30.256
Congestive Heart Failure	6.999	6.888	6.791	6.751	6.751
Acute Myocardial Infarction	9.715	9.553	9.443	9.441	9.442
Unstable Angina and Other Acute Ischemic Heart Disease	6.438	6.331	6.260	6.262	6.262
Heart Infection/Inflammation, Except Rheumatic	16.113	15.984	15.888	15.866	15.866
Hypoplastic Left Heart Syndrome and Other Severe Congenital Heart Disorders	6.323	6.111	5.905	5.794	5.792
Major Congenital Heart/Circulatory Disorders	1.778	1.651	1.493	1.391	1.389
Atrial and Ventricular Septal Defects, Patent Ductus Arteriosus, and Other Congenital Heart/Circulatory Disorders	1.202	1.090	0.952	0.872	0.871
Specified Heart Arrhythmias	4.399	4.213	4.049	3.984	3.983
Intracranial Hemorrhage	15.936	15.685	15.510	15.504	15.504
Ischemic or Unspecified Stroke	8.574	8.456	8.381	8.396	8.396
Cerebral Aneurysm and Arteriovenous Malformation	3.865	3.650	3.490	3.433	3.432
Hemiplegia/Hemiparesis	4.815	4.703	4.625	4.610	4.610
Monoplegia, Other Paralytic Syndromes	3.627	3.487	3.391	3.361	3.361
Atherosclerosis of the Extremities with Ulceration or Gangrene	15.571	15.296	15.096	15.012	15.011

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Vascular Disease with Complications	18.826	18.672	18.564	18.569	18.569
Pulmonary Embolism and Deep Vein Thrombosis	15.291	15.130	15.023	15.041	15.042
Lung Transplant Status/Complications	30.468	30.333	30.245	30.256	30.256
Cystic Fibrosis	20.415	19.976	19.647	19.686	19.687
Chronic Obstructive Pulmonary Disease, Including Bronchiectasis	0.435	0.348	0.231	0.149	0.147
Asthma	0.435	0.348	0.231	0.149	0.147
Fibrosis of Lung and Other Lung Disorders	4.116	3.973	3.845	3.789	3.788
Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections	10.256	10.199	10.157	10.177	10.177
Kidney Transplant Status	16.425	16.083	15.843	15.848	15.848
End Stage Renal Disease	39.805	39.631	39.521	39.592	39.593
Chronic Kidney Disease, Stage 5	7.087	6.923	6.771	6.675	6.673
Chronic Kidney Disease, Severe (Stage 4)	7.087	6.923	6.771	6.675	6.673
Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism	1.126	0.939	0.750	0.559	0.555
Miscarriage with Complications	1.126	0.939	0.750	0.559	0.555
Miscarriage with No or Minor Complications	1.126	0.939	0.750	0.559	0.555
Completed Pregnancy With Major Complications	3.159	2.712	2.427	2.240	2.240
Completed Pregnancy With Complications	3.159	2.712	2.427	2.240	2.240
Completed Pregnancy with No or Minor Complications	3.159	2.712	2.427	2.240	2.240
Chronic Ulcer of Skin, Except Pressure	1.941	1.836	1.731	1.675	1.675
Hip Fractures and Pathological Vertebral or Humerus Fractures	5.725	5.450	5.215	5.124	5.123

Factor	Platinum	Gold	Silver	Bronze	Catastrophic
Pathological Fractures, Except of Vertebrae, Hip, or Humerus	1.574	1.428	1.264	1.147	1.145
Stem Cell, Including Bone Marrow, Transplant Status/Complications	30.468	30.333	30.245	30.256	30.256
Artificial Openings for Feeding or Elimination	14.575	14.480	14.443	14.551	14.553
Amputation Status, Lower Limb/Amputation Complications	8.195	7.923	7.727	7.631	7.630

Table A.3:
Infant Risk Adjustment Models Factors

Group	Platinum	Gold	Silver	Bronze	Catastrophic
Extremely Immature * Severity Level 5 (Highest)	378.927	377.561	376.491	376.507	376.508
Extremely Immature * Severity Level 4	194.401	193.057	192.003	191.981	191.981
Extremely Immature * Severity Level 3	46.419	45.304	44.390	44.236	44.234
Extremely Immature * Severity Level 2	46.419	45.304	44.390	44.236	44.234
Extremely Immature * Severity Level 1 (Lowest)	46.419	45.304	44.390	44.236	44.234
Immature * Severity Level 5 (Highest)	190.323	189.030	188.013	188.027	188.028
Immature * Severity Level 4	85.852	84.500	83.442	83.437	83.437
Immature * Severity Level 3	46.419	45.304	44.390	44.236	44.234
Immature * Severity Level 2	28.986	27.832	26.907	26.738	26.736
Immature * Severity Level 1 (Lowest)	28.986	27.832	26.907	26.738	26.736
Premature/Multiples * Severity Level 5 (Highest)	156.158	154.846	153.824	153.791	153.791

Group	Platinum	Gold	Silver	Bronze	Catastrophic
Premature/Multiples * Severity Level 4	32.573	31.292	30.290	30.173	30.173
Premature/Multiples * Severity Level 3	17.215	16.169	15.315	15.020	15.016
Premature/Multiples * Severity Level 2	8.942	8.081	7.334	6.884	6.876
Premature/Multiples * Severity Level 1 (Lowest)	6.222	5.557	4.867	4.376	4.367
Term * Severity Level 5 (Highest)	130.728	129.499	128.518	128.414	128.413
Term * Severity Level 4	16.874	15.867	15.038	14.685	14.681
Term * Severity Level 3	6.324	5.648	4.969	4.448	4.438
Term * Severity Level 2	3.857	3.319	2.700	2.139	2.128
Term * Severity Level 1 (Lowest)	1.639	1.321	0.772	0.358	0.350
Age1 * Severity Level 5 (Highest)	54.166	53.499	52.963	52.894	52.892
Age1 * Severity Level 4	9.298	8.787	8.351	8.169	8.167
Age1 * Severity Level 3	3.380	3.034	2.676	2.465	2.461
Age1 * Severity Level 2	2.155	1.873	1.549	1.320	1.316
Age1 * Severity Level 1 (Lowest)	0.572	0.441	0.274	0.199	0.197
Age 0 Male	0.685	0.637	0.608	0.554	0.553
Age 1 Male	0.145	0.127	0.106	0.081	0.081

Appendix B: Prescription Drug Modeling Econometric Model and Constraints

This paper describes two main ways in which prescription drug utilization can convey information about the health status of health plan enrollees: an imputation approach and a severity approach. In designing a model to predict individual expenditures, these two approaches require different specifications within a linear econometric model. In fact, both approaches can be implemented as special cases of a single, generalized model by imposing constraints on the coefficient estimates.

In the prescription drug modeling section, we refer to a "flexible" model (which we call "generalized" in this appendix) that allows three independent effects of a drug-diagnosis pair on predicted expenditures. Three indicator variables – the HCC, the RXC, and an HCC×RXC interaction term – enter the model specification as follows:

$$E = \alpha_0 + \beta_1 HCC + \beta_2 RXC + \beta_3 HCC \times RXC$$

In this generalized version of the model, there are three different levels of predicted incremental expenditure relevant to this drug-diagnosis pair:

HCC only	β_1
RXC only	β_2
HCC and RXC	$\beta_1 + \beta_2 + \beta_3$

This is the most flexible specification possible in a linear model.

To impose the severity approach within the structure of this model, we can impose the constraint $\beta_2 = 0$. Now the model will predict incremental expenditures for someone with the HCC only (β_1), and a different level of expenditures for someone with the HCC and the RXC ($\beta_1 + \beta_3$). Having only the RXC without the associated HCC, would have no impact on predicted expenditures.

The imputation approach can also be achieved with coefficient constraints. By imposing the restriction $\beta_1 = \beta_2 = -\beta_3$, the model will predict the same incremental expenditures for all three cases:

HCC only	β_1
RXC only	β_1
HCC and RXC	β_1

This "pure" imputation approach is probably unrealistic, because any drug class that can impute the presence of a diagnosis almost certainly contains information about the severity and costs associated with that diagnosis, even if only because taking the drug itself entails additional costs beyond what someone with only the diagnosis incurs.

Additional variations are also possible. For example, another model we present in this chapter, the "drug dominant" model, can be obtained by rest restricting $-\beta_3 = \beta_1 > 0$ and $\beta_2 \geq \beta_1$. This predicts two levels of incremental expenditure: β_1 for people with the HCC only, and β_2 for people with the RXC (regardless of whether they have the HCC or not).

This generalized linear econometric model can therefore accommodate several different types of relationships between drug classes and diagnoses. The key question then becomes which specification and constraints the model should use. These decisions will depend on broader criteria including predictive accuracy, incentives for prescription drug utilization and diagnosis coding, and clinical face validity of the models.

Appendix C: List of RXCs

Table C.1:
List of RXCs

RXC	RXC Label	Payment
4.01	Antihistamine	
4.02	First Generation Antihistamines and Phenothiazine Derivatives	
8.02	Severe Bacterial Infections	Y
8.03	HIV	Y
8.04	Interferons	Y
8.05	Polyenes	Y
8.07	Other Anti-Infectives	
8.08	Quinolones	Y
8.09	Miscellaneous Antimycobacterials	Y
8.10	Antituberculosis Agents	Y
8.11	Hep C Antivirals	Y
10.01	Antineoplastic Agents	Y
12.01	Cholinergic Agents	
12.02	Anticholinergic Agents	Y
12.03	Alpha Adrenergic Agonists	Y
12.04	Alpha-Adrenergic Blocking Agents	
12.05	Beta-Adrenergic Blocking Agents	
12.06	Muscle Relaxants, High Severity	Y
12.07	Muscle Relaxants, Low Severity	
12.08	Other/Miscellaneous Autonomic Drugs	
12.09	Beta Adrenergic Agonists	
20.01	Heparins	Y

RXC	RXC Label	Payment
20.02	Anticoagulent except Heparins	Y
20.03	Platelet-Aggregation Inhibitors	Y
20.04	Hematopoietic Agents	Y
20.05	Hemostatics	Y
20.06	Blood Formation, Coagulation, Thrombosis, Except 20.01-20.05	
24.01	Class IB and Class III Antiarrhythmics	Y
24.02	Antiarrhythmics, except Class IB and Class III Antiarrhythmics	Y
24.03	Antilipemic	
24.04	Anti-Hypertension	
24.05	Central Alpha Agonists	
24.06	High Severity Diuretic	Y
24.07	Erectile Dysfunction	
24.08	Miscellaneous Cardiac	Y
24.09	Diuretics	
24.10	Direct Vasodilators	
24.11	Nitrates/Nitrites	Y
24.12	Miscellaneous Vasodilating	
24.13	Cardiotonic Agents	Y
26.01	Cellular Therapy	
28.01	Anesthetics	
28.02	Opiate agonists	Y
28.03	Analgesics, other/low severity	
28.04	Anticonvulsants: Hydantoins, Oxazolidinediones, and Succinimides	Y
28.05	Other Anticonvulsants	Y
28.06	Antipsychotics - high severity	Y
28.07	Antipsychotics - low severity	Y
28.08	Antidepressants	
28.09	Amphetamines	
28.10	Misc anorexigenic agents and stimulants	
28.11	Barbiturates	
28.12	Benzodiazepines	
28.13	Misc. Anxiolytics, sedatives & hypnotics	
28.14	Antimanic agents	Y
28.15	Antimigraine agents	
28.16	Antiparkinsonian agents	Y
28.18	Misc CNS agents	
28.19	Opiate partial agonists	
32.01	Contraceptives	
36.01	Diagnostic agents	
40.01	Acidifying Agent	Y

RXC	RXC Label	Payment
40.02	Other Electrolytic	Y
40.03	ESRD	Y
40.04	Severe Kidney Disease	Y
40.05	Hospital Only Category (IV Only)	
40.06	Ammonia Detoxicants	Y
40.07	Medical Supplies	
44.01	Enzymes	Y
48.01	Respiratory agents, low severity	
48.02	Orally Inhaled Preparations and Leukotriene Modifiers	
48.03	Mucolytic agents	Y
48.04	Misc respiratory tract agents	Y
48.05	Cystic Fibrosis tranmembrane conductance regulator agents	Y
48.06	Antifibrotic respiratory agents	Y
48.07	Phosphodiesterase Type 4 Inhibitors (severe COPD)	Y
52.01	Ocular NSAIDs	Y
52.02	Glaucoma	Y
52.04	Mydriatics	
52.05	Other EENT	
52.06	Ocular Surgery	
56.01	Cholelitholytic Agents	Y
56.02	HT3 Receptor Antagonists and Other Antiemetic Agents	
56.03	Proton-Pump Inhibitors, Histamine H2-Antagonists, and Protectants	
56.04	Anti-Inflammatory Agents Used to Treat Inflammatory Bowel Disease	Y
56.05	Prostaglandins	
56.06	Antihistamines for Gastrointestinal	
56.07	Digestants	Y
56.08	Prokinetic Agents	
56.09	Other Gastrointestinal Drugs, Except 56.01-56.08	
64.01	Heavy Metal Antagonists	Y
68.01	Adrenals	
68.02	Androgens	
68.03	Other Hormone	
68.04	Estrogen Agonist	
68.05	Gonadotropins	
68.06	Insulins	Y
68.07	Antidiabetic Agent	Y
68.08	Antihypoglycemic Agents	Y
68.09	Parathyroid	Y
68.10	Pituitary	Y
68.11	Somatropin Agonists and Antagnoists	

RXC	RXC Label	Payment
68.12	Thyroid Agents	
68.13	Antithyroid Agents	
68.14	Somatostatin Agonists and Antagonists	
72.01	Local Anesthetics	
76.01	Oxytocics	
78.01	Radioactive Agents	
80.01	Serums	Y
80.02	Toxoids and Vaccines	
84.01	Skin and Mucous Membrane Agents	
86.01	Genitourinary Smooth Muscle Relaxants	
86.02	Respiratory Smooth Muscle Relaxants	
88.01	Vitamin K Activity	Y
88.02	Vitamins, Except Vitamin K Activity	
92.01	Alcohol Deterrents	Y
92.02	BPH	
92.03	Antidotes	Y
92.04	Antigout	Y
92.05	Biologic Response Modifiers Acting on the Central Nervous System	Y
92.06	Rheumatoid Arthritis	Y
92.07	Gonadotropin-Releasing Hormone	Y
92.08	Transplant Antirejection Immunosuppressive Agents	Y
94.01	Devices	
96.01	Pharmaceutical Aids	
99.01	Miscellaneous	
99.02	Excluded Therapeutic Classes	