DEPARTMENT OF HEALTH & HUMAN SERVICES
Centers for Medicare & Medicaid Services
Center for Consumer Information and Insurance Oversight
200 Independence Avenue SW
Washington, DC 20201

HHS-Operated Risk Adjustment Technical Paper on Possible Model Changes

October 26, 2021

Contents

Overview

The Patient Protection and Affordable Care Act (ACA)[1] 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 lowest-risk enrollees and avoid the highest-risk enrollees. The risk adjustment program is intended to achieve this goal by mitigating the effect of risk selection on premiums by transferring funds from issuers that enroll lower-than-average risk populations to issuers that enroll higher-than-average risk populations.

The Department of Health and Human Services (HHS) has actively sought comment and received feedback on the federal risk adjustment requirements from the beginning of the program's development.[2] This formal and informal feedback has been instrumental in developing program requirements, as well as identifying areas for further study and potential refinements for future benefit years. Most recently, HHS proposed, but did not finalize, several updates to the HHS risk adjustment models in the 2022 Payment Notice.[3] In the 2022 Payment Notice, in response to comments, HHS reiterated its commitment to continue to consider potential changes that could increase the current models' predictive accuracy and acknowledged stakeholders' desire for additional analyses and information on the proposed model specification updates.[4] This paper provides additional detail and analyses on the model updates proposed in the 2022 Payment Notice. It also includes information on HHS' ongoing evaluation of the state payment transfer formula's current cost-sharing reduction induced demand factors (CSR IDFs).

Chapter 1 provides an explanation of how the risk adjustment models currently work, an overview of previous model updates, and an explanation of the current focus on improving the models' predictive accuracy for certain subpopulations. Chapter 2 explores ways to improve the current models' predictive accuracy for the lowest-risk enrollees. Chapter 3 explores ways to improve the predictive accuracy of the enrollment duration factors included in the models. Chapter 4 explores ways to improve the current models' predictive accuracy for the highest-risk

---

[1] The Patient Protection and Affordable Care Act (Pub. L. 111-148), was enacted on March 23, 2010. The Health Care and Education Reconciliation Act of 2010 (Pub. L. 111-152), which amended and revised several provisions of the ACA, was enacted on March 30, 2010. In this paper, we refer to the two statutes collectively as the Affordable Care Act or ACA.

[2] See, e.g., Planning and Establishment of State-Level Exchanges; Request for Comments Regarding Exchange-Related Provisions in Title I of the Patient Protection and Affordable Care Act; Request for comments, 75 FR 45584 (August 3, 2010); and the March 2016 Risk Adjustment Methodology White Paper (March 24, 2016), available at: https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf.

[3] See Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2022 and Pharmacy Benefit Standards; Updates to State Innovation Waiver (Section 1332 Waiver) Implementing Regulations; Proposed Rule, 85 FR 78572 at 78581 – 78586 (December 4, 2020). Also see Patient Protection and Affordable Care Act; HHS Notice of Benefit and Payment Parameters for 2022 and Pharmacy Benefit Standards; Updates to State Innovation Waiver (Section 1332 Waiver) Implementing Regulations; Final Rule, 86 FR 24140 at 24151 - 24162 (May 5, 2021).

[4] See 86 FR at 24151 – 24162.

enrollees. Chapter 5 provides the predictive ratios and R-squared statistics for all of the model specification changes examined in this paper and discusses the estimated impact of those changes. Additionally, Chapter 5 includes information on how we are planning to simulate the impact of the risk adjustment modeling changes discussed in Chapters 2, 3 and 4 in this paper on plan liability risk scores (PLRS) and transfers under the state payment transfer formula. Finally, Appendix A shares information on the current CSR IDFs in the state payment transfer formula, stakeholder feedback on these factors, and options HHS is evaluating to improve prediction of CSR enrollees' plan liability.[5] We are accepting comments on this paper (including the Appendices) at RARIPAYMENTOPERATIONS@cms.hhs.gov with the subject line of "2021 Model Update Technical Paper" until November 26, 2021.

Chapter 1: Overview and Current Status of Risk Adjustment Models

This chapter provides an introduction to the HHS-operated risk adjustment program and a high-level summary of the state payment transfer formula.[6] The chapter begins by providing background on the history and purpose of the risk adjustment program and the principles that have guided the development of the current HHS risk adjustment models. It also provides details on the calculation of the risk scores and transfer amounts under the state payment transfer formula. Finally, the chapter concludes with a summary of previous changes to the risk adjustment models since the beginning of the HHS risk adjustment program and identifies several areas that could further improve the current models' predictive accuracy for certain subpopulations, which are addressed in more detail in later chapters of this paper.

1.1 Overview of Risk Adjustment

1.1.1 History

Section 1343 of the ACA[7] established a permanent risk adjustment program in which states collect charges from health insurance issuers that enroll lower-than-average risk populations and provide payments to health insurance issuers that enroll higher-than-average-risk populations, such as those with chronic conditions, thereby reducing incentives for issuers to avoid higher-risk enrollees. Consistent with section 1321(c)(1) of the ACA, HHS is responsible for operating the risk adjustment program on behalf of any state that does not elect to do so.[8] Since the 2017 benefit year, HHS has operated the risk adjustment program for all 50 states and

[5] As explained in Appendix A, the data and analyses in Chapters 2 – 5 and the forthcoming transfer simulation will not reflect any changes to the CSR IDFs. Although we did not propose any changes to the CSR IDFs in the 2022 Payment Notice, we wanted to share information with stakeholders about HHS' ongoing evaluation of these factors.
[6] The state payment transfer formula refers to the part of the HHS risk adjustment methodology that calculates payments and charges at the state market risk pool level prior to the calculation of the high-cost risk pool payment and charge terms that apply beginning with the 2018 benefit year. See, e.g., 81 FR at 94080.
[7] 42 U.S.C. § 18063.
[8] 42 U.S.C. § 18041(c).

the District of Columbia.[9] The risk adjustment program applies to risk adjustment covered plans in the individual and small group (including merged) markets, inside and outside the Exchanges.[10] Since the 2014 benefit year, the risk adjustment program has transferred funds from issuers with plans that insure lower-than-average risk enrollees to plans that insure higher-than-average risk enrollees, resulting in billions of dollars being transferred among issuers annually. In 2020 for example, the absolute value of transfers was $11.17 billion, or 7.5 percent of premium, between issuers in all risk pools (excluding the high-cost risk pool).[11]

As described in the 2014 Notice of Benefit and Payment Parameters (2014 Payment Notice) final rule, the risk adjustment methodology developed by HHS is based on the premise that premiums should reflect the differences in plan benefits, quality, and efficiency—not the health status of the enrolled population.[12] The state payment transfer formula that is part of the HHS methodology determines each risk adjustment covered plan's risk score and state transfer payment or charge amount based on the actuarial risk of enrollees, the actuarial value (AV) of coverage, the cost of doing business in local rating areas accounting for care utilization, and the effect of different cost-sharing levels on utilization. Thus, the HHS risk adjustment methodology predicts average group costs to account for risk across plans, in keeping with the Actuarial Standards Board's Actuarial Standards of Practice for risk classification.[13]

Risk adjustment payments and charges under the HHS-operated program are budget neutral, meaning the total amount of risk adjustment charges collected from issuers are equal to the total amount of risk adjustment payments made. These budget-neutral transfers reduce the incentive for issuers to avoid the highest-risk enrollees and reduce the influence of risk selection on the premiums that plans charge.[14]

1.1.2 Purpose and Goals of Risk Adjustment

The purpose of the risk adjustment program is to reduce the influence of risk selection on plan premiums as well as to reduce the incentive for plans to avoid enrolling higher-than-average risk enrollees. Without risk adjustment, plans that enroll a higher proportion of high-than-average-risk enrollees would need to charge a higher average premium (across all of their enrollees) to be financially viable. The intent of risk adjustment is to allow a plan enrolling a higher proportion of high-than-average-risk enrollees to charge the same average premium (other

[9] Massachusetts ran its own risk adjustment program for benefit years 2014-2016. See page 5 of the March 2016 Risk Adjustment Methodology White Paper (March 24, 2016), available at: https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf.

[10] See 45 CFR 153.20 for definition of risk adjustment covered plans.

[11] See the Summary Report on Permanent Risk Adjustment Transfers for the 2020 Benefit Year at: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/RA-Report-BY2020.pdf.

[12] 78 FR 15409 at 15417.

[13] See the Actuarial Standards Board's Actuarial Standards of Practice for risk classification at: http://www.actuarialstandardsboard.org/asops/risk-classification-practice-areas/.

[14] For a further discussion of the budget neutral approach for the HHS-operated risk adjustment program, see the 2020 Payment Notice final rule, 84 FR at 17480 – 17482.

factors being equal) as a plan enrolling a higher proportion of lower-than-average-risk enrollees, shifting the focus of plan competition to plan benefits, quality, efficiency, and value.

1.2 The Current HHS Risk Adjustment Models

1.2.1. Principles of Risk Adjustment

To determine a plan's actuarial risk, the HHS risk adjustment models use an enrollee's demographics and chronic health status information 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). Hierarchical condition categories (HCCs) are a critical component of accurate risk prediction as they dictate which diagnostic codes are included in the models, how they are grouped, and how those groupings interact. There are ten principles of risk adjustment which guide the HHS risk adjustment models' diagnostic classification system.[15]

<u>Principle 1:</u> Diagnostic categories should be clinically meaningful.
<u>Principle 2:</u> Diagnostic categories should predict medical (including drug) expenditures.
<u>Principle 3:</u> Diagnostic categories that will affect payments should have adequate sample sizes to permit accurate and stable estimates of expenditures.
<u>Principle 4:</u> 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.
<u>Principle 5:</u> The diagnostic classification should encourage specific coding.
<u>Principle 6:</u> The diagnostic classification should not reward coding proliferation.
<u>Principle 7:</u> Providers should not be penalized for recording additional diagnoses (monotonicity).
<u>Principle 8:</u> The classification system should be internally consistent (transitive).
<u>Principle 9:</u> The diagnostic classification should assign all ICD-10-CM codes (exhaustive classification).
<u>Principle 10:</u> Discretionary diagnostic categories should be excluded from payment models.

Consistent with these risk adjustment principles, the HHS risk adjustment models exclude 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 models also exclude diagnoses that do not add significantly to costs (e.g., non-melanoma forms of skin cancer). 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 were used to make tradeoffs among principles where they conflicted. Some examples of conflict among

---

[15] A detailed explanation of the ten principles of risk adjustment can be found in the March 2016 Risk Adjustment Methodology White Paper (March 24,2016), available at: https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf. Also see Potential Updates to HHS–HCCs for the HHS-operated Risk Adjustment Program (June 17, 2019) paper, available at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf.

principles include tradeoffs between clinical meaningfulness (principle 1) and adequate sample sizes (principle 3), as well as between encouraging specific coding (principle 5) and predictive accuracy (principle 2).

In addition to the ten principles, the current HHS risk adjustment program has two main process components: (1) a method for measuring risk selection by calculating risk scores,[16] and (2) a method for quantifying financial impacts by calculating risk adjustment transfers at the state market risk pool level.[17] The following sections detail the formulas used to calculate risk scores and transfer amounts under the state payment transfer formula.[18]

1.2.2 Calculation of the Risk Score

To determine the risk score for each risk adjustment covered plan, HHS developed separate models for adults, children, and infants to account for clinical and cost differences in each age group. These HHS risk adjustment models predict plan liability for an average enrollee based on that person's age, sex, and diagnoses (HCCs), producing a risk score.

The risk score for an enrollee is defined as the total predicted relative plan liability expenditures for the enrollee based on the relevant HHS risk adjustment model for the enrollee's age group (adult, child, or infant) and plan metal level. For the metal level of the enrollee's plan, the total predicted relative plan liability expenditures, or individual enrollee risk scores, are calculated based on the following factors: an age factor, a demographic factor, diagnostic (or HCC) factor(s), an enrollment duration factor (adult models only), prescription drug (or RXC) factor(s) (adult models only), and a CSR factor (if applicable). The age model factor assigns risk scores to each of nine age categories which range from 21 to 64 in five-year increments in the adult models and is a significant driver of health risk differences. For a child enrollee, the relative risk assigned to an individual is the sum of the applicable age, demographic, and diagnostic factor(s); and for infants, it is the sum of the appropriate maturity/disease severity category and the male additive term.[19] If applicable, a multiplicative adjustment for induced utilization is made to the risk score for enrollees in individual market CSR plan variations offered through the Exchanges.

Enrollees' model-calculated risk scores are then weighted by enrollment months and aggregated to arrive at a plan-level average risk score for that plan's enrollees within a rating area. As previously stated, the enrollment-weighted average risk score of all enrollees in a particular risk adjustment covered plan (also referred to as the plan liability risk score or PLRS) within a geographic rating area is one of the inputs into the state payment transfer formula, which determines the state transfer payment or charge that an issuer will receive or be required to pay for that plan in a given benefit year for the applicable state market risk pool.

---

[16] See The HHS-HCC Risk Adjustment Model for Individual and Small Group Markets under the Affordable Care Act, 2014. Available at: https://www.cms.gov/mmrr/Downloads/MMRR2014_004_03_a03.pdf

[17] See Risk Transfer Formula for Individual and Small Group Markets Under the Affordable Care Act, 2014. Available at: https://www.cms.gov/mmrr/Downloads/MMRR2014_004_03_a04.pdf.

[18] See supra note 6.

[19] In the infant models, there are two additive terms for sex, for age 0 males and age 1 males, which account for higher morbidity and infant mortality in the male infant population (females are the reference group for the mutually-exclusive categories).

The PLRS is a crucial component for calculating risk adjustment transfer payments or charges under the state payment transfer formula. The transfer formula also incorporates several other factors into its calculation of transfers, as described in the next section.

1.2.3 Calculations Under the State Payment Transfer Formula

As explained above, the state payment transfer formula averages all individual risk scores in risk adjustment covered plans in a state market risk pool, and uses the PLRS, combined with other factors, to calculate the payment and charge amounts at the state market risk pool level.

The first step in calculating transfers under the state payment transfer formula is finding the T_{PMPM}, or the transfer amount per billable member per month (PMPM) for plan i, and is described below.

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

The transfer formula may be understood to be composed of two key higher-level terms, both enclosed in the square brackets: the risk term ($\frac{PLRS_i \cdot IDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot IDF_i \cdot GCF_i)}$) and the rating term ($\frac{AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i}{\sum (s_i \cdot AV_i \cdot ARF_i \cdot IDF_i \cdot GCF_i)}$).

The first of these, the risk term, defines the revenue required by a plan (relative to the statewide market average). It is determined by three component variables:

1. The PLRS, which reflects plan i's AV as well as the plan's enrollee health status risk;[20]
2. An induced demand factor (IDF), which reflects the anticipated induced demand associated with plan i's cost-sharing (metal) level; and
3. A geographic cost factor (GCF), which accounts for differences in premium due to allowable geographic rating variation.

The second term, the rating term, defines the revenue that a plan can be expected to generate given the allowable rating factors (relative to the statewide market average). It is determined by four component variables:

1. AV, which adjusts for relative differences between the plan actuarial value in a market;
2. An Allowable Rating Factor (ARF), which accounts for the impact of allowable rating factors (age or family tier) based on state rating method;
3. An IDF; and
4. A GCF.

The denominators of the risk and rating terms in the transfer equation express statewide average required revenue and allowable premium, respectively. The statewide average required revenue and allowable premium include the same component variables from the numerator multiplied by each plan's share of statewide enrollment, or s_i. Dividing these terms by the

[20] The cost-sharing reduction induced demand factor (CSR IDF) is applied, if applicable, at the enrollee's enrollment-period level prior to calculation of the plan-level $PLRS_i$ and reflects the anticipated induced demand associated with lower cost sharing (see Appendix A for further discussion of CSR IDFs).

respective statewide average expresses the plan's revenue requirement and allowable revenue relative to the average plan offered in the state. Transfers are calculated by converting the applicable factors into dollar amounts by multiplying them by the statewide enrollment-weighted market average plan premium ($\bar{P}_s$ in the transfer formula). 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 = 1.75).

The final transfer formula is as follows.

$$T_i = T_{PMPM} * \sum_b M_b$$

Structured as shown above, the final transfer formula calculates T_i, plan i's total transfer amount for a rating area. The total transfer amount for each plan for the applicable state market risk pool is calculated by multiplying PMPM transfers (T_{PMPM}), the payment or (if negative) charge to plan i for each member month of enrollment, by the plan's total billable member months ($\sum_b M_b$), where M_b is the number of months during the risk adjustment period the billable enrollee b is enrolled in plan i.

1.3 Previous Risk Adjustment Model and Transfer Formula Updates

HHS has continuously analyzed and refined the HHS risk adjustment models and transfer formula through notice-and-comment rulemaking and has shared detailed analysis in prior white papers. [21] Past updates to the models and transfer formula are described below.

In the 2014 Payment Notice,[22] we established the first set of HHS risk adjustment models that were used for the 2014 and 2015 benefit years. These initial models were developed using the Truven Health Analytics 2010 MarketScan® Commercial Claims and Encounters database (MarketScan® data).

In the 2016 Notice of Benefit and Payment Parameters (2016 Payment Notice),[23] we finalized the recalibration of the HHS risk adjustment models using blended 2011, 2012, and 2013 MarketScan® data to develop final risk adjustment factors for the 2016 benefit year. The goal of this update to the models was to provide risk adjustment factors that reflected more recent treatment patterns and costs while simultaneously providing greater stability within the

[21] See 2011 Risk Adjustment Implementation Issues, September 12, 2011, available at: https://www.cms.gov/CCIIO/Resources/Files/Downloads/riskadjustment_whitepaper_web.pdf and March 2016 Risk Adjustment White Paper, March 24, 2016, available at: https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf and Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program, June 17, 2019, available at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf.

[22] 78 FR 15409.

[23] 80 FR 10749.

risk adjustment program and minimizing risk score volatility by blending factors from three years of data.

In the 2017 Notice of Benefit and Payment Parameters (2017 Payment Notice),[24] we finalized the introduction of preventive services into the simulation of plan liability as part of the recalibration of the risk adjustment models beginning with the 2017 benefit year. We also finalized the use of blended 2012, 2013 and 2014 MarketScan® data to recalibrate the risk adjustment model factors for the 2017 benefit year.

In the 2018 Notice of Benefit and Payment Parameters (2018 Payment Notice),[25] we finalized several model updates:

1. Adding enrollment duration factors: We finalized the addition of enrollment duration factors to the adult models starting with the 2017 benefit year. The enrollment duration factors are used in the calculation of adult enrollee risk scores to account for additional risk associated with enrollees with partial-year enrollment. They do so through a set of 11 enrollment duration binary indicatory variables that signify that an enrollee had exactly one to 11 months of enrollment in a given plan. The model-estimated risk value of these indicators decreases monotonically from one to 11 months, reflecting the increased annualized costs associated with fewer months of enrollment.
2. Adding prescription drug factors: We finalized the addition of prescription drugs to the adult models by creating indicators for categories of prescription drugs that interact with HCCs, or RXC-HCC interactions, beginning with the 2018 benefit year. The purpose of including RXCs in the adult risk adjustment models is to improve prediction by identifying health conditions that may otherwise be underreported, as well as to detect diagnoses that may not be present in medical claims data because a patient with a long-term chronic condition has not visited a provider for that condition during their plan enrollment. The presence of a specific drug is used to indicate that the associated diagnosis is possibly present and is captured by the RXC factor, while the RXC-HCC interaction captures differing levels of severity of illness that treatment with a specific drug may indicate. There are currently 10 RXC-HCC interactions used in the adult models.[26]
3. Establishing a high-cost risk pool: We finalized the establishment of a high-cost risk pool beginning with the 2018 benefit year to account for the risk of high-cost enrollees and help ensure that transfers under the state payment transfer formula better reflect typical actuarial risk without the influence of extremely high cost conditions. To account for this high-cost risk pool during recalibration, we exclude a percentage of costs above a certain threshold level (currently the parameters are set at a $1 million threshold and a

[24] 81 FR 12203.
[25] 81 FR 94058.
[26] In the 2018 Payment Notice (81 FR 94058), we finalized the addition of prescription drugs to the adult risk adjustment models by creating indicators for 12 categories of prescription drugs, beginning with the 2018 benefit year. In the 2019 Payment Notice (83 FR 16930), we finalized the removal of the two severity-only RXCs (RXC 11: Ammonia Detoxicants, and RXC 12: Diuretics, Loop and Select Potassium-Sparing).

coinsurance rate of 60 percent[27]) in the calculation of enrollee-level plan liability risk scores. This exclusion ensures that the risk adjustment factors account for risk associated with HCCs and RXCs excluding those extreme costs, because the average risk associated with HCCs and RXCs is better accounted for without inclusion of the high-cost enrollees.

4. Reducing the statewide average premium: We finalized a 14 percent reduction to the statewide average premium in the state payment transfer formula, beginning with the 2018 benefit year, to reflect the portion of administrative costs that does not vary by claims.

In the 2018 Payment Notice, we also finalized the proposal to collect and use enrollee-level External Data Gathering Environment (EDGE) data for recalibration of the risk adjustment models in future benefit years. Because MarketScan® data generally reflects the large group market, the use of enrollee-level EDGE data allows for the annual recalibration of the risk adjustment models to better reflect the populations in the individual and small group (or merged) markets that are subject to risk adjustment. "Enrollee-level EDGE dataset" in this paper refers to the national dataset used for risk adjustment model recalibration. This dataset contains limited information generated by reports HHS receives from issuers' EDGE servers, and excludes plan, state, and other information necessary for calculating risk adjustment transfers. We also finalized the use of blended 2013, 2014 and 2015 MarketScan® data to recalibrate the risk adjustment model factors for the 2018 benefit year.

In the 2019 Notice of Benefit and Payment Parameters (2019 Payment Notice),[28] we finalized the use of equally blended factors from 2014 MarketScan®, 2015 MarketScan®, and the 2016 enrollee-level EDGE dataset. We also finalized the removal of two severity-only RXCs (RXC 11: Ammonia Detoxicants, and RXC 12: Diuretics, Loop and Select Potassium-Sparing) from the adult models for the 2019 benefit year and beyond.

In the 2020 Notice of Benefit and Payment Parameters (2020 Payment Notice),[29] we finalized a pricing adjustment for the plan liability simulation for the Hepatitis C RXC to more closely reflect the expected average additional plan liability of the Hepatitis C RXC for the 2020 benefit year. We also finalized the use of equally blended factors from separately solved 2015 MarketScan®, 2016 enrollee-level EDGE, and 2017 enrollee-level EDGE datasets to recalibrate the risk adjustment model factors for the 2020 benefit year.

In the 2021 Notice of Benefit and Payment Parameters (2021 Payment Notice),[30] we finalized the proposed updates to the risk adjustment models' HCCs to transition the HHS-HCC clinical classifications to the International Classification of Diseases, 10th Revision (ICD-10) diagnosis codes to reflect more recent diagnosis code information and claims data.[31] To develop

---

[27] See, e.g., 84 FR at 17466 through 17468 and 86 FR at 24183 through 24186.
[28] 83 FR 16930.
[29] 84 FR 17454.
[30] 85 FR 29164
[31] The initial HHS-HCC clinical classification, in place from the 2014 through the 2020 benefit year, was based on International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes. In June 2019, we released a paper on the potential updates to the HHS-HCC clinical classification to incorporate ICD-10-

the HHS-HCC classification changes, we used 2016 and 2017 benefit years' enrollee-level EDGE claims data, which reflected the first two full years of ICD-10 diagnosis coding on claims. To guide the reclassification process, we used the 10 principles that guided the creation of the original HHS-HCC diagnostic classification system as finalized in the 2014 Payment Notice[32] and enumerated in section 1.2.1 of this paper. In the 2021 Payment Notice,[33] we also finalized the use of 2016, 2017 and 2018 enrollee-level EDGE datasets to recalibrate the risk adjustment model factors for the 2021 benefit year and adopted a similar recalibration approach (to use the three most recent years of enrollee-level EDGE datasets available) for future benefit year recalibrations, unless changed in rulemaking. Additionally, we continued to make a pricing adjustment for the plan liability simulation for the Hepatitis C RXC.

In the 2022 Notice of Benefit and Payment Parameters (2022 Payment Notice),[34] we finalized the use of 2016, 2017 and 2018 enrollee-level EDGE datasets to recalibrate the risk adjustment model factors for the 2022 benefit year (the same years used to recalibrate the 2021 models), and finalized a policy for the 2022 benefit year and beyond to publish final risk adjustment coefficients using the three most recent years of EDGE data that are available when we publish the proposed Payment Notice for the applicable benefit year. We also continued to make a pricing adjustment for the plan liability simulation for the Hepatitis C RXC.

In the 2022 Payment Notice, we did not finalize several proposed model specification updates that were intended to improve the models' predictive accuracy for certain subpopulations, including incorporation of a two-stage weighted specification in the adult and child models, revisions to the adult models' enrollment duration factors, and the addition of new severity and transplant indicators interacted with HCC count factors in the adult and child models.[35] We discuss these 2022 Payment Notice proposals in greater detail in Chapters 2-5 of this document.

1.4 Predictive Accuracy of the Current HHS Risk Adjustment Models

The predictive accuracy of a risk adjustment model is typically evaluated using predictive ratios (PRs), calculated as the ratio of predicted to actual weighted mean plan liability expenditures. The predictive ratio represents how well the model has done on average at predicting plan liability for that subpopulation. If prediction is perfect, mean predicted expenditures will equal mean actual expenditures, and the PR will be 1.00. Throughout this

CM diagnosis codes. See Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program, June 17, 2019. Available at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf.

[32] 78 FR 15409.

[33] 85 FR 29164.

[34] 86 FR 24140.

[35] On July 16, 2021 CMS released "Updated 2022 Benefit Year Final HHS Risk Adjustment Model Coefficients". In part 2 of the HHS Notice of Benefit and Payment Parameters for 2022 final rule (2022 Payment Notice), we finalized the risk adjustment models for the 2022 benefit year, but the 2022 benefit year coefficients included in this final rule contained a few errors that impacted certain factors for the adult models. Therefore, consistent with 45 C.F.R. § 153.320(b)(1)(i), we announced the final 2022 benefit year final risk adjustment adult model coefficients that included some minor revisions: https://www.cms.gov/files/document/updated-2022-benefit-year-final-hhs-risk-adjustment-model-coefficients-clean-version-508.pdf.

paper, we present and discuss PRs from our risk score models. The predicted plan liability for an enrollee is their sum of model factors (i.e., risk score). We calculate separate risk scores for each metal level, reflecting differences in actuarial value.[36]

Predictive accuracy is also measured by the R-squared statistic. This statistic reflects the predictive accuracy of the model for variations in individual expenditures, whereas PRs are used to test the performance of the models among subgroups (for example, enrollees in certain age or expenditure categories, and enrollees with and without HCCs). HHS uses both PRs and R-squared statistics to evaluate the HHS risk adjustment models.

One of the benefits of the current HHS risk adjustment models is that they perfectly predicted mean plan liability expenditure by age-sex factor, HCC or HCC group, RXC, and enrollment duration factor (PRs for all factors included in the model equal exactly 1.00 in the enrollee-level EDGE datasets used for model recalibration).[37,38] The PRs of 1.00 result from a regression that estimates the models to accurately predict mean expenditures for all factors in the regression (i.e., the mean prediction for all variables in the model will equal the actual mean). This perfect prediction can be seen in Figure 1.1.

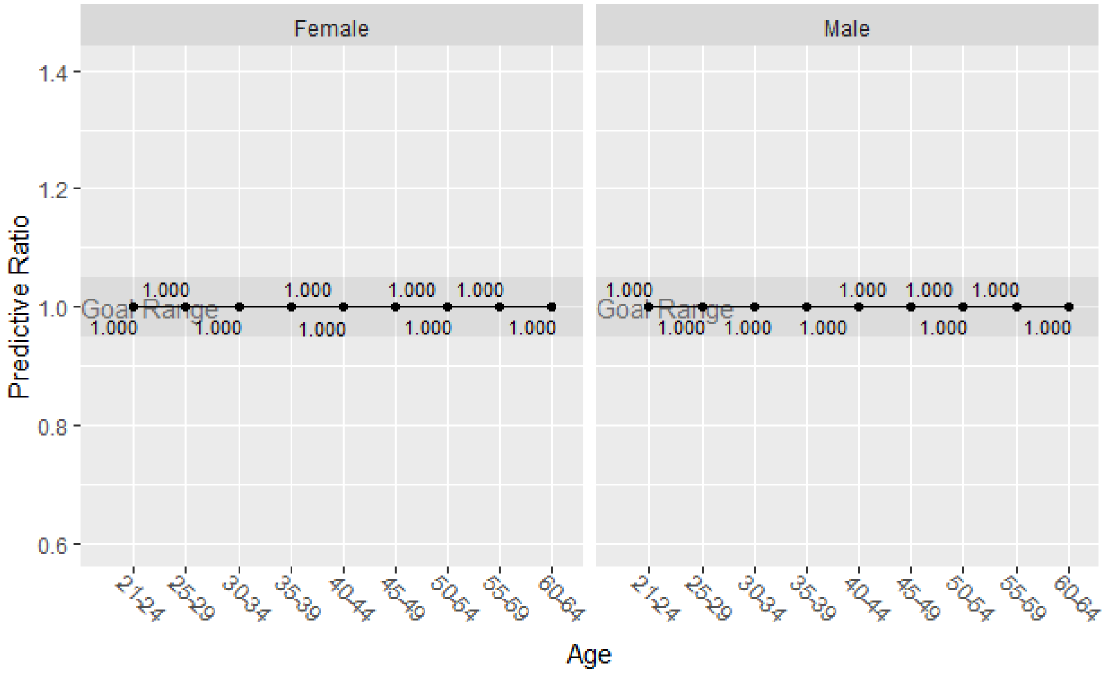
---

[36] In Appendix A, we explore PRs that explicitly address differences in induced demand and other factors by metal and CSR, which are included in the risk term of the state payment transfer formula. In particular, some of the PRs in Appendix A include the metal and CSR IDFs. We refer to these as "risk term PRs," since they evaluate the full risk term of the transfer formula.

[37] PRs are also 1 or very close to 1 in validation samples. The PRs calculated in this paper are calculated using the same samples on which the models were calibrated. However, as is common practice in evaluating model fit, we also tested splitting the sample for calibration and validation purposes and the results were unchanged. Further, for purposes of the analysis in this paper, we calculated PRs for at least three data years. The results always appear the same, so we generally only report results in this paper from the 2018 data year, which was the most recently available dataset at the time that we ran these analyses in preparation for announcing the proposals in the proposed 2022 Payment Notice.

[38] There are two exceptions involving RXC variables where the coefficients are constrained such that the PRs are less than 1.0.

Figure 1.1: Adult Silver Plan Model Predictive Ratios by Age-Sex, 2018 Enrollee-Level EDGE Dataset



Despite the predictive accuracy for mean plan liability expenditure in the national calibration sample, some stakeholders have expressed concern that the current risk adjustment models underpredict plan liability for the subpopulation of the lowest-risk enrollees. These stakeholders recommended that CMS address this potential underprediction to prevent a misalignment of incentives that would discourage enrollment of lower risk enrollees relative to higher risk ones and negatively impact the overall health of the risk pool. CMS has continued to consider whether improvements could be made to the predictive accuracy of the risk adjustment models for certain subgroups.[39]

Figure 1.2 shows the predictive ratios for the current adult[40] silver model by decile of mean predicted plan liability expenditure (or "predicted risk") using the 2018 enrollee-level EDGE dataset. We use the silver models in tables and figures throughout this paper, as the silver models apply to the most enrollees, and we generally see similar patterns in the models for the other metal levels. We also focus on the adult models in tables and figures through this paper because the adult models apply to the vast majority of enrollees. Decile 1 represents the lowest decile of predicted risk, 0-10 percent, and decile 10 represents the highest decile of predicted

[39] See, e.g., the 2021 Payment Notice proposed rule, 85 FR at 7101 - 7104 and the 2022 Payment Notice proposed rule, 85 FR at 78582 – 78586.

[40] "Adult enrollees" refers to all adults in the enrollee-level EDGE dataset used to recalibrate the risk adjustment models, which includes adults ages 21-64 without any capitated claims. Adults with capitated claims are excluded from the dataset, as the risk adjustment models are used to evaluate enrollees' expenditures, and capitated claims do not provide meaningful and comparable cost (allowed charges) data in comparison to non-capitated claims for recalibration purposes. We also were concerned that methods for computing and reporting derived amounts from capitated claims could be inconsistent across issuers and would not result in reliable data for recalibration or the analyses in this paper.

risk, 90-100 percent, which we have further segmented into the top 5 percent, 1 percent, and 0.1 percent of predicted expenditures. As seen in Figure 1.2, the current adult models underpredicted plan liability for the lowest-risk enrollees (deciles 1 through 6), overpredicted plan liability for some medium-risk enrollees (decile 8), and underpredicted risk for the highest-risk enrollees (the top 0.1 percent).

Figure 1.2: Adult Silver Plan Model Predictive Ratios by Decile, 2018 Enrollee-Level EDGE Dataset

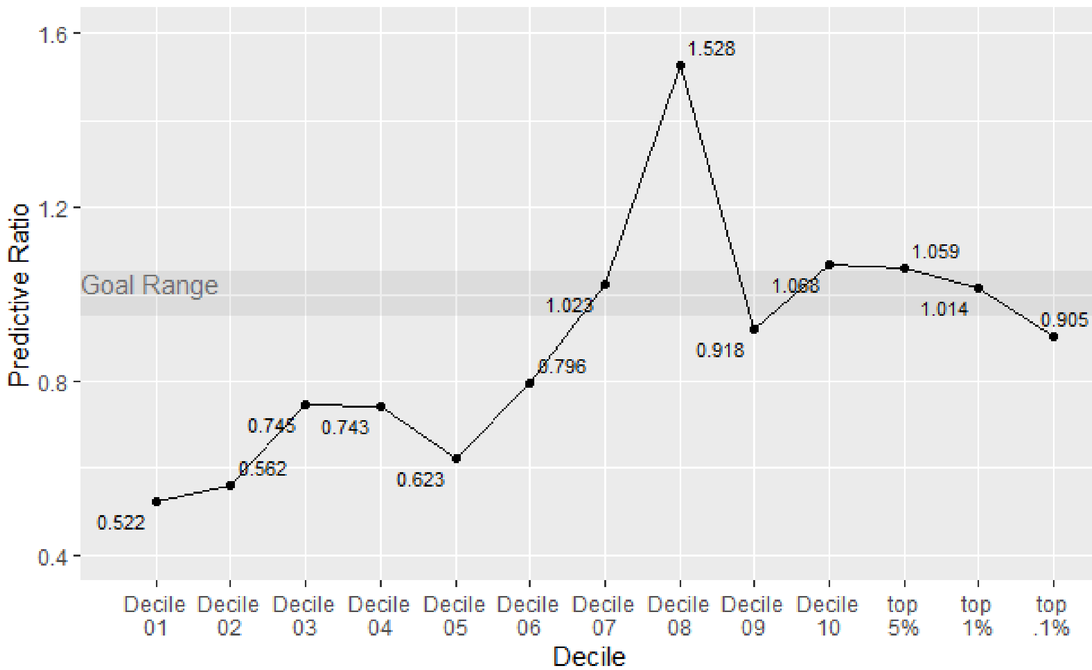


Figure 1.3 shows the risk for enrollees without HCCs in the current adult models was underpredicted, the risk for enrollees with 1 to 5 HCCs tended to be overpredicted, and the risk for enrollees with 7 or more HCCs was underpredicted.

Figure 1.3: Adult Silver Plan Model Predictive Ratios by Number of HCCs, 2018 Enrollee-Level EDGE Dataset

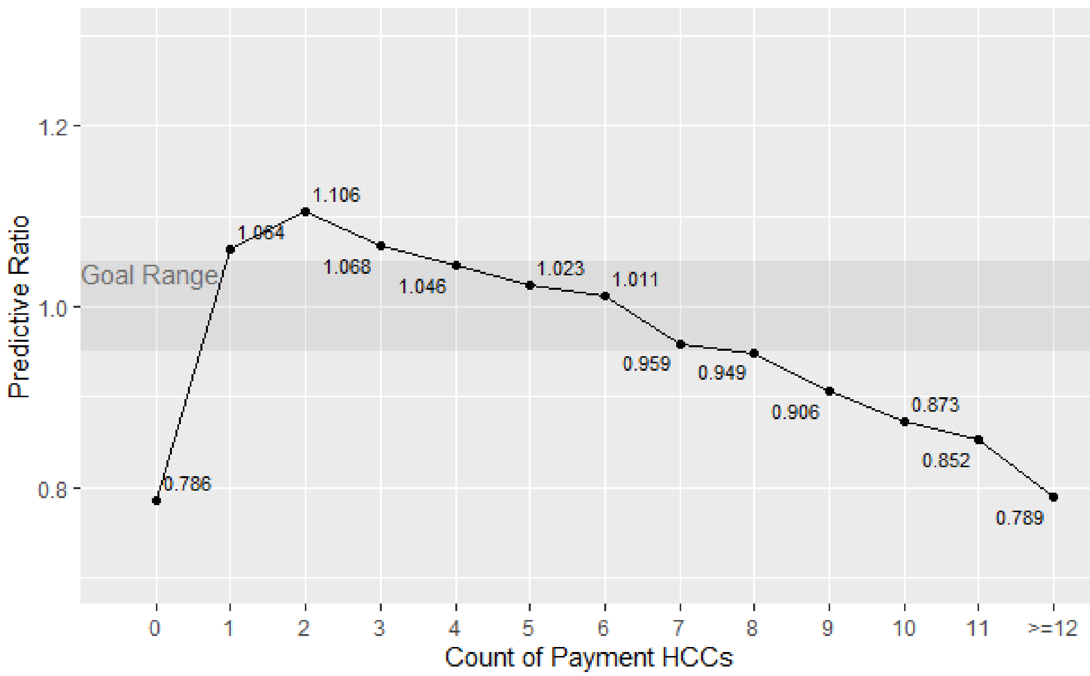
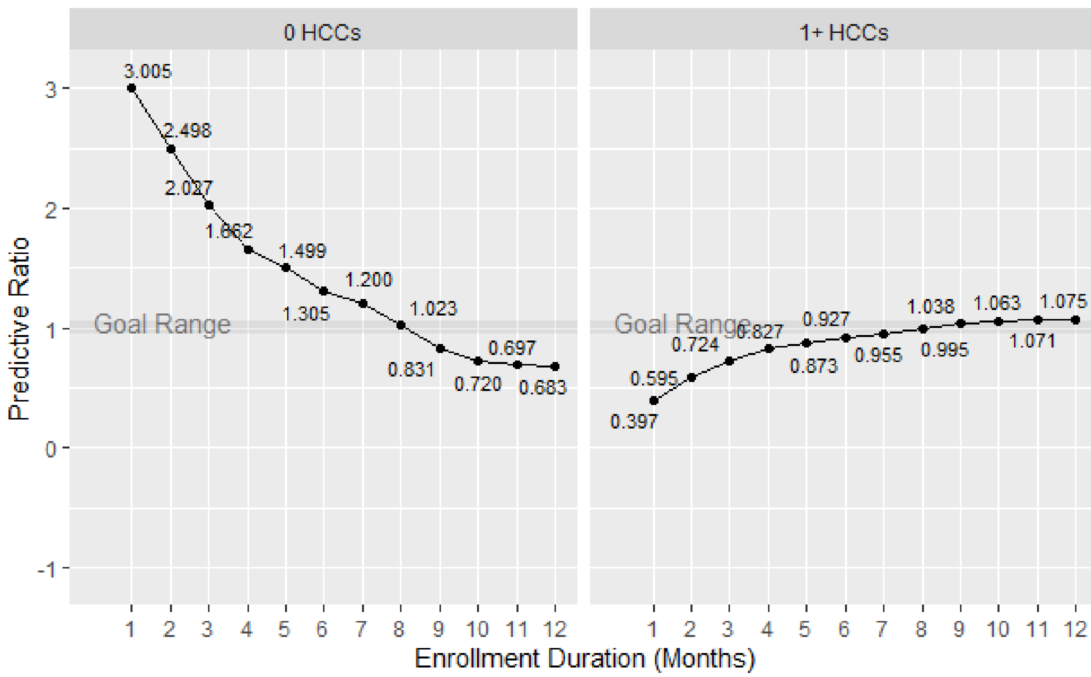


Figure 1.4 shows that the models overpredicted silver plan liability for adult enrollees without HCCs with 1 to 7 months of enrollment and underpredicted silver plan liability for adult enrollees with HCCs with 1 to 6 months of enrollment in the 2018 enrollee-level EDGE dataset. The underprediction among adults without HCCs with 9-12 months of enrollment duration, shown in figure 1.4, is consistent with the broader underprediction of full year enrollees without HCCs, which we discuss in greater detail in Chapter 2.[41]

Figure 1.4 Adult Silver Plan Model Predictive Ratios by Enrollment Duration and Number of HCCs, 2018 Enrollee-Level EDGE Dataset



Although the current models perfectly predicted mean plan liability expenditure in the enrollee-level EDGE data used for model recalibration, continued study and analysis has identified three subpopulations for which the predictive accuracy of the current HHS risk adjustment models could be improved:

1. *Lowest-Risk Enrollees*. The current models underpredicted plan liability for the lowest-risk enrollees (that is, enrollees in low-risk deciles and enrollees without HCCs). The underprediction of the lowest-risk enrollees can be seen in the PRs in the HHS risk adjustment adult models by decile of predicted costs (Figure 1.2) and by number of payment HCCs that the enrollee has (Figure 1.3), using the 2018 enrollee-level EDGE dataset at the silver metal level. As seen in Figure 1.2, the adult models underpredicted the lowest-risk enrollees (risk decile 1) by over 50 percent. The prediction improved mostly monotonically with enrollee risk, with the exception of a substantial

[41] Although 8 months of enrollment is overpredicted, it falls within the goal range of +/- 5 percent (e.g., .95 to 1.05), with a PR of 1.023. Similarly, although 7 and 8 months of enrollment are underpredicted, they fall within the goal range with PRs of .955 and .995 respectively.

overprediction in risk decile 8. As seen in Figure 1.3, the enrollees without HCCs in the adult models were underpredicted, enrollees with 1 to 5 HCCs tended to be slightly overpredicted, whereas enrollees with 7 or more HCCs tended to be somewhat underpredicted. This trend can be seen across all metal levels to varying degrees. Since approximately 80 percent of enrollees in the individual and small group (or merged) markets do not have HCCs, this underprediction applies to a large portion of the risk adjustment population.[42] We did further analysis and found the same patterns, including the underprediction of the enrollees without HCCs, in the child models.[43] Although the model underprediction of lowest-risk enrollees may appear large in percentage terms, it is not large in dollar terms because the lowest-risk enrollees are not costly.

2. *Partial-Year Adult Enrollees*. The current models overpredicted plan liability for partial-year adult enrollees without HCCs and underpredicted plan liability for partial-year (1 to 6 months of enrollment) adult enrollees with HCCs. As described in the proposed 2021 Payment Notice[44] and the proposed 2022 Payment Notice,[45] our analysis of the 2017 and 2018 enrollee-level EDGE datasets found that the current enrollment duration factors are driven by enrollees with HCCs. That is, partial-year enrollees with HCCs had higher PMPM expenditures on average compared to full-year enrollees with HCCs. On the other hand, partial-year enrollees without HCCs did not have substantively different PMPM expenditures compared to full-year enrollees without HCCs.[46]

   We also found that the current factors masked the much stronger relationship between enrollment duration and costs that exists for the small fraction of partial-year enrollees with HCCs. To further illustrate this relationship, Table 1.1 shows the PMPM allowed charges by enrollment duration and by presence of HCCs for adult enrollees, using the 2018 enrollee-level EDGE dataset. Among adults with HCCs, we found that average monthly costs were approximately 3.5 times *higher* (a difference of $3,877 monthly) for adults enrolled for 1 month compared to adults enrolled for 12 months. Among adults

---

[42] See the Summary Report on Permanent Risk Adjustment Transfer for the 2019 Benefit Year (Figure 3) at: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/RA-Report-BY2019.pdf.
[43] The infant models utilize 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; 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. Our analysis found that the infant models more accurately predicted costs for the lowest-risk infants. The underprediction trends for the lowest-risk adults and children were not observed for infants.
[44] 85 FR 29164 at 29189.
[45] 85 FR 78572 at 78585-78586.
[46] We also explained in the proposed 2021 Payment Notice that we found partial year enrollees in the child models did not have the same risk differences as partial year enrollees in the adult models and they tended to have similar risk to their full year counterparts. See 85 FR at 7103 and 7104. In the infant models, we found that partial year enrollees had higher expenditures on average compared to full-year enrollees; however, the incorporation of enrollment duration factors created interaction issues with the current severity and maturity factors in the infant models and did not have a meaningful impact on the general predictive accuracy of the infant models. Ibid.

without HCCs, however, we found that average monthly costs were 26 percent *lower* (a difference of $46 monthly) for adults enrolled for 1 month compared to adults enrolled for 12 months. Monthly allowed charges generally decreased monotonically with enrollment duration among adult enrollees with HCCs and increased monotonically with enrollment duration among adult enrollees without HCCs. The differences were much more pronounced for shorter enrollment durations (e.g., 1-2 months) than longer enrollment durations (e.g., 11-12 months). Similar relationships held for adults with other partial-year enrollment durations.

Table 1.1: PMPM Allowed Charges by Enrollment Duration and Presence of HCCs (Adult Enrollees), 2018 Enrollee-Level EDGE Dataset

	Adults with HCCs		Adults without HCCs	
Enrollment Duration	**Number of Enrollees (% of all)**	**PMPM Allowed Charges**	**Number of Enrollees**	**PMPM Allowed Charges**
1 month	41,877 (4%)	$5,452	1,051,103	$129
2 months	66,665 (7%)	$3,452	892,311	$132
3 months	86,321 (10%)	$2,753	780,634	$146
4 months	106,445 (12%)	$2,287	764,035	$145
5 months	102,282 (14%)	$2,146	628,808	$152
6 months	106,744 (16%)	$1,994	577,431	$151
7 months	102,756 (17%)	$1,950	501,700	$152
8 months	108,911 (18%)	$1,838	489,152	$154
9 months	104,911 (19%)	$1,760	440,063	$155
10 months	114,432 (20%)	$1,687	453,519	$155
11 months	148,023 (21%)	$1,610	559,719	$159
12 months	2,551,057 (23%)	$1,575	8,583,428	$175

Note: The percentages of number of adult enrollees are row percentages. For example, 4 percent of all enrollees with 1 month of enrollment had an HCC.

As seen in Table 1.1 above, the fraction of partial-year enrollees who had HCCs is small, compared to the approximately 20 percent of the overall population who had HCCs. As compared to the 23 percent of full-year enrollees who had HCCs, only 10 percent of the approximately 5.2 million adults in the 2018 enrollee-level EDGE dataset enrolled for 1 to 6 months had HCCs. However, enrollees with 1 to 6-month enrollment duration factors that had HCCs had PMPM costs that were much higher than full-year enrollees with HCCs. The difference in PMPM costs between enrollees with 1 to 6-month enrollment duration factors who had HCCs and their counterparts without HCCs was even more pronounced. The result is that the current enrollment duration factors, which are estimated using data for partial-year enrollees as a whole, with and without HCCs, reflect an inverse relationship between enrollment duration and costs, resulting in factors fairly close to zero. We further discuss the prediction issues for partial year enrollees, and our consideration of options to improve prediction for this subpopulation, in Chapter 3.

3. *Very Highest-Risk Enrollees.* The current models also underpredicted plan liability for the very highest-risk enrollees (that is, those in the top 0.1 percent risk percentile and those enrollees with the most HCCs). As seen in Figure 1.2 above, the current models underpredicted adult silver plan enrollees' plan liability in the top 0.1 risk percentile by 9 percent. Additionally, as shown in Figure 1.3, enrollees with more than 6 HCCs were underpredicted in the current adult models and, in particular, enrollees with 10 or more HCCs were underpredicted by approximately 17 percent. As discussed above, the models' underprediction of lowest-risk enrollees is not large in dollar terms, because the lowest-risk enrollees are not costly. However, this is not the case for enrollees in the highest deciles of risk. Enrollees in risk decile 10 represent roughly 74.29 percent of actual plan liability, compared to only 1.36 percent for enrollees in risk decile 1. The most expensive enrollees tended to have severe acute illness HCCs. These enrollees were often hospitalized, received ICU care, and were frequently among individuals exceeding the $1 million high-cost risk pool claim threshold.[47] For example, we found that 50 percent of enrollees reaching the $1 million claims threshold have HCC 2 (Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock). These enrollees also tended to have a large number of HCCs that have a non-linear effect on costs, meaning that the actual cost for these enrollees was higher than simply adding the incremental effects of each HCC together. We further discuss the underprediction for the very highest-risk enrollees, and our consideration of options to improve prediction for this subpopulation, in Chapter 4.

1.4.1 Purpose of this Paper

Following our study and consideration of these issues and ways to improve the current models' predictive accuracy for the subpopulations described above, we proposed several model specification changes in the proposed 2022 Payment Notice.[48] We proposed updating the HHS risk adjustment models by adding a two-stage weighted model specification to the adult and child models to weight lower-risk enrollees more heavily, modifying the enrollment duration factors in the adult models to better capture the increased costs of partial-year adult enrollees with HCCs, and adding severity and transplant indicators interacted with HCC counts factors to the adult and child models to better predict costs of the very highest-risk enrollees. After consideration of comments, we did not finalize these proposed model specification updates. However, we reiterated our commitment to continue to consider potential changes to increase the current models' predictive accuracy for these subpopulations and acknowledged stakeholders' desire for additional analyses and information on the proposed model specification updates.[49] In

[47] In the 2018 Payment Notice, we incorporated a high-cost risk pool into the HHS risk adjustment methodology. To account for this high-cost risk pool during recalibration, we exclude a percentage of costs above a certain threshold level (as of the 2022 benefit year, the parameters are set at a $1 million threshold and a coinsurance rate of 60 percent) in the calculation of enrollee-level plan liability risk scores. This ensures that the risk adjustment factors account for risk associated with HCCs and RXCs excluding those extreme costs, because the average risk associated with HCCs and RXCs is better accounted for without inclusion of the high-cost enrollees.
[48] 85 FR 78572 at 78583-86.
[49] 86 FR at 24151 – 24162.

this paper, we describe our detailed analyses of the changes proposed but not finalized in the 2022 Payment Notice.[50]

For the vast majority of the analysis in this paper, we used the 2018 benefit year enrollee-level EDGE dataset to test the model specification changes as this dataset was the most recently available dataset at the time that we ran these analyses in preparation for announcing the proposals in the proposed 2022 Payment Notice. In certain cases, we also used the 2016 and 2017 benefit year enrollee-level EDGE datasets when we were testing model specification changes across data years. Additionally, in limited cases, we used the 2019 benefit year enrollee-level EDGE dataset to run analysis in this paper, such as some testing of the enrollment duration factors, as this dataset has recently become available. Throughout this paper, we prioritized using recent data that would allow us to avoid repeatedly re-running analyses as new datasets became available.

We believe that the information in this paper will assist stakeholders in better understanding these changes as our intention is to re-propose these changes in future rulemaking. Finally, we also share information in Appendix A on the current CSR IDFs in the state payment transfer formula, stakeholder feedback over the years on these factors, and options HHS is evaluating to improve prediction of CSR enrollees' plan liability.[51]

[50] 85 FR 78572 at 78583-86.

[51] As explained in Appendix A, the data and analyses in Chapters 2 – 5 and the forthcoming transfer simulation will not reflect any changes to the CSR IDFs. While we did not propose any changes to the CSR IDFs in the 2022 Payment Notice, we wanted to share information with stakeholders about HHS' ongoing evaluation of these factors. We may propose changes to these factors for future benefit years and would seek stakeholder feedback on any such changes in a future rulemaking.

Chapter 2: Improving the Models' Predictive Accuracy for Lowest-Risk Enrollees – Two-Stage Weighted Approach

This chapter focuses on the underprediction of risk for the lowest-risk enrollee subpopulations – enrollees without HCCs[52] and low-cost enrollees[53] – in the current HHS risk adjustment models. This chapter also discusses our consideration of options to improve the models' predictive accuracy for these subpopulations. Specifically, this chapter explores our considerations that informed the proposed two-stage weighted approach in the proposed 2022 Payment Notice, and provides analyses on the impact of the two-stage weighted approach.

2.1 Current Models' Prediction of Costs for Lowest-Risk Enrollees

As described in Chapter 1, risk adjustment uses models to predict costs based on certain identifying characteristics of a population. In the HHS risk adjustment models, enrollee risk scores are determined by age, sex, health conditions (as indicated by certain diagnoses and, for adult enrollees, prescription drugs), and enrollment duration (for adult enrollees), and are used to predict health expenditures for each metal level. This modeling process works well for an average person within each level of each risk adjustment model factor cell. To the degree a risk adjustment factor's population costs widely vary within the same factor, and plan enrollee costs deviate significantly from the population average, the factor may not precisely predict the health risk across all plan enrollees.

Risk scores for enrollees without HCCs are mainly influenced by the applicable age, sex and enrollment duration factors in the current HHS risk adjustment models.[54] These age, sex and enrollment duration (if applicable) factors serve two functions:

(1) They represent the prediction of plan liability for the approximately 80 percent of enrollees without HCCs ("lowest-risk"); and

(2) They represent a portion of the prediction of plan liability for the approximately 20 percent of enrollees with HCCs ("highest-risk"), with the remaining portion consisting of the liability associated with health conditions.

For all coefficients, we use least squares estimates of the HHS risk adjustment models that balance the fit in these two populations. This constraint affects the age, sex and enrollment duration (if applicable) factors. In particular, looking at enrollees without HCCs separately, the least squares estimates of the age-sex factors would mean that these factors would almost entirely predict plan liability for these enrollees. Considering the highest-risk enrollees separately, the least squares estimates would jointly optimize the age-sex factors and the highest-risk enrollees' HCC coefficients to achieve the best fit. For the highest-risk enrollees, this results

[52] When we refer to the enrollees without HCCs, we are referring to enrollees without payment HCCs throughout this paper.

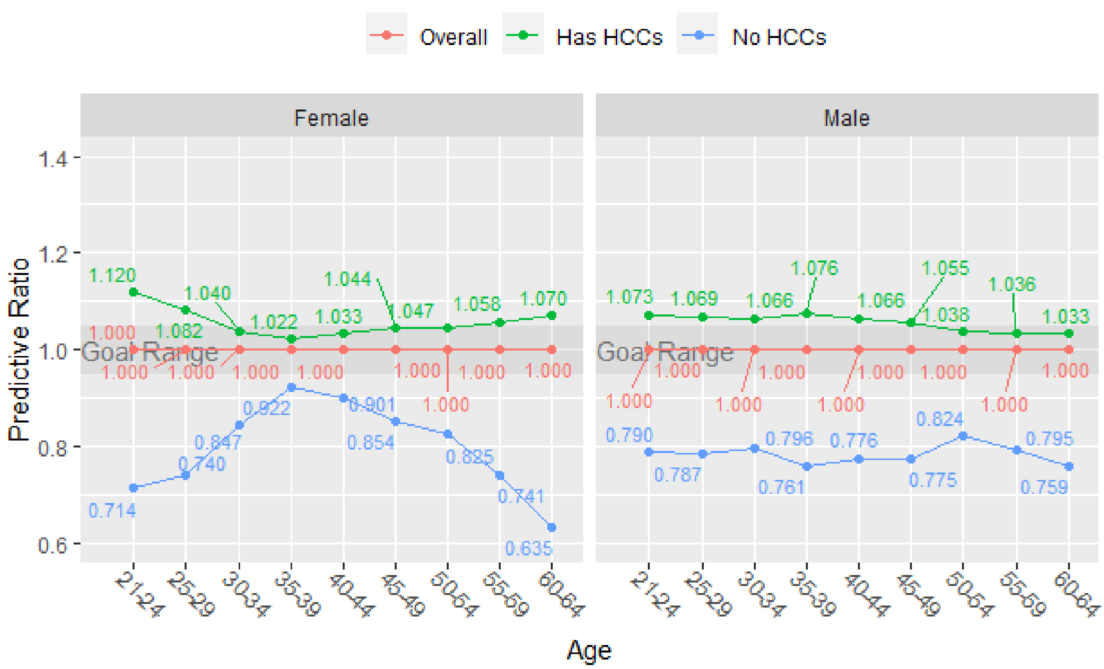
[53] We define low-cost enrollees based on enrollees in the lower deciles of predicted plan liability, which is different than considering enrollees based on HCCs. Assessing the impact of model specification changes based on payment HCCs would yield a different result than assessing the impact of model specification changes based on prediction deciles because of the presence of outliers in prediction deciles.

[54] Adult enrollees without HCCs can have RXC factors, but this is uncommon. Outside of the regression model, adult, child and infant enrollees may also have a CSR adjustment factor.

in age-sex factor estimates that are very small and are less than mean plan liability by age-sex (and may even be negative). The age-sex factor estimates the average of the higher plan liability estimates for the lowest-risk enrollees with the lower plan liability estimates for the highest-risk enrollees, resulting in a combined estimate for the entire population. Therefore, because plan liability for the lowest-risk enrollees (e.g. enrollees without HCCs and low-cost enrollees) is almost entirely comprised of age-sex factors that are averaged across a population that includes the highest-risk enrollees, the average plan liability of the lowest-risk enrollees is underpredicted by the HHS risk adjustment models.

To illustrate how model predictions of plan liability by age-sex factor work, we calculated three sets of PRs for the age-sex factors. First, we calculated PRs for each age-sex factor for all enrollees; second, we calculated PRs for each age-sex factor for enrollees without HCCs; third, we calculated PRs for each age-sex factor for enrollees with HCCs. Figure 2.1 shows that the current adult models underpredict the risk by age-sex category for enrollees without HCCs, overpredicts the risk by age-sex category for enrollees with HCCs (1+ HCCs), but accurately predicts average risk overall by age-sex category when combining both enrollees with and without HCCs.

Figure 2.1: Adult Silver Plan Model Predictive Ratios by Age-Sex, 2018 Enrollee-Level EDGE Dataset



By addressing the underprediction of costs associated with lowest-risk enrollees (enrollees without HCCs and low-cost enrollees) in the risk adjustment models, we expect to encourage retention and entry into the individual and small group (including merged) markets by plans that enroll a higher proportion of this subpopulation of enrollees. These plans are at greater risk of exiting the market if risk adjustment payments or charges undercompensate for the true plan liability of the lowest-risk enrollees. By taking steps to improve the models' ability to predict the true costs for these enrollees, we hope to improve incentives for issuers to create plans that encourage these individuals to enroll in coverage and improve the individual and small group (including merged) market risk pools. Therefore, in the proposed 2022 Payment Notice, we proposed to update the adult and child models by adding a two-stage weighted model specification to weight lowest-risk enrollees more heavily. The discussion that follows in Section 2.2 outlines options we considered to improve the current models' predictive accuracy for the lowest-risk enrollees. Section 2.3 provides analyses on the impact of the two-stage weighted approach.

2.2 Improving Prediction for Lowest-Risk Enrollees – Two-Stage Weighted Model Specification

Since 2016, we considered a variety of different options to address the underprediction of plan liability for lowest-risk enrollees in various Payment Notices[55] and shared our concerns about the limitations of many of these options. For example, in the 2018 Payment Notice, we explored the creation of two separate age-sex factors in the adult models for enrollees with and without HCCs to address the models' underprediction of costs for the lowest-risk enrollees.[56] However, upon further analysis, we found that since the estimated "high-risk" age-sex factors (meaning the age-sex factors for enrollees with HCCs) contribute substantially less to the overall risk coefficient than do the estimated "low-risk" age-sex factors (meaning the age-sex factors for enrollees without HCCs), the sum of the "high-risk" age-sex factor and the HCC factor(s) for a "high-risk" enrollee could be less than if that enrollee had instead been given just the "low-risk" age-sex factor. In other words, the predicted cost for a high-risk enrollee could be lower than the predicted cost for a low-risk enrollee of the same age and sex. Or, put another way, predicted cost could fall for a low-risk no-HCC enrollee when the enrollee acquires an HCC diagnosis. This lacks face validity and violates the principles of risk adjustment by penalizing issuers for reporting additional conditions.

We also considered adding a non-linear term to the adult and child models to allow greater flexibility in our model specification and improve the fit of our model to non-linear data.[57] While the non-linear approach appeared to at least partially address the underprediction of the lowest-risk enrollee subpopulations, it would result in very different model factor estimates and involved somewhat complex transformations of coefficients to obtain estimates of the marginal effects of the HCC on costs. In addition, the need to explain and justify a fundamentally different model structure is a significant disadvantage, especially when compared to the simplicity of the standard linear regression framework used for the current model framework, which is widely understood and accepted, making it easier for issuers to use in rate setting. Further, the non-linear modeling approach faced operational challenges in achieving consistent convergence of iterative estimates.

[55] See 81 FR 94058 at 94082-94084, 85 FR 7088 at 7101-7104, and 85 FR 29164 at 29188-29190.
[56] 81 FR 94058 at 94083.
[57] 85 FR 29164 at 29188- 29190. We also discussed using a non-linear and full counts model in the Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program paper, available at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf.

In addition to the non-linear approach, we investigated an HCC counts approach similar to that implemented in the Medicare Advantage risk adjustment models beginning with Calendar Year 2020.[58] While adoption of a full HCC counts approach for the HHS risk adjustment models for the individual and small group (including merged) markets would improve prediction for the very highest-risk enrollees, it raised concerns of gaming wherein issuers may be incentivized to code additional HCCs to increase the enrollee's HCC count and thus increase their risk adjustment payment (or reduce their charge). Also, similar to the non-linear approach, we found that the full HCC counts approach would result in very large model coefficient changes. These concerns led us to not prefer the full HCC count approach in the policy context of the HHS risk adjustment models.[59]

After consideration of the aforementioned approaches, we wanted to consider an alternative that would limit the number of changes to the current models to promote stability, mitigate gaming concerns, and minimize issuer burden with regard to rate setting. As a result, we focused on weighting the calibration sample observations in model estimation to prioritize improved accuracy of predictions for the lowest-risk enrollee subpopulations. The process that we considered consisted of two stages. The first stage involved estimating the current models to predict plan liability, then using those predicted values to create a set of weights calculated as the inverse of predicted plan liability, such that observations with lower predicted plan liability had higher weights. These weights were then used in a second stage regression to weight each observation to re-estimate predicted values of plan liability, outputting the coefficients for the final HHS risk adjustment models. We refer to this process as the "two-stage weighted approach."[60]

The conceptual reasoning for the two-stage weighted approach is to retain the simple, linear, additive structure of the current models while forcing the models to better predict costs for lowest-risk enrollees. The two-stage weighted approach would address the HHS risk adjustment models' underprediction for this subpopulation and thereby improve model calibration across deciles of predicted expenditures. Recognizing that the reciprocal weighting by predicted expenditures is not a usual or standard procedure in risk adjustment,[61] we are supplementing the explanation of the two-stage weighted approach in the proposed 2022 Payment Notice with this Technical Paper to explain our analysis of the two-stage weighted approach and its impact on the risk adjustment models and transfers calculated under the state payment transfer formula. As detailed in the proposed 2022 Payment Notice and outlined further below, our analysis found the gains in model performance outweigh the potential drawbacks.

---

[58] See pages 22-26 of the Report to Congress: Risk Adjustment in Medicare Advantage (December, 2018) available at: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf. Also see: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Advance2020Part1.pdf.

[59] As we discuss in Chapter 4, we found that an interacted HCC counts model specification mitigates these gaming concerns and improves prediction compared to the current model, including for the lowest-risk enrollees (enrollees without HCCs and low-cost enrollees).

[60] Although there are two stages in the proposed model, the two-stage weighted model should not be confused with two-stage least squares (2SLS) techniques used to analyze structural equations.

[61] For example, Medicare Advantage does not use a two-stage weighting in its model recalibration.

To develop this approach, we considered a variety of weights for the HHS risk adjustment model estimation. Because all of our risk markers (regressors) are binary variables, but a weight variable needs to be continuous, it was natural to consider functions of the aggregation of the risk markers into predicted cost as regression weights. Since we want to reweight the higher risk individuals with higher predicted cost, we considered the reciprocal of functions of predicted cost as weights. We also considered the reciprocal of predicted cost raised to powers ranging from zero (equal weighting) to 2, in increments of 0.5. This series included the reciprocal of the square root of predicted cost (exponent = 0.5), the reciprocal of predicted cost (exponent = 1.0), the reciprocal of the predicted cost times the reciprocal of the square root of predicted cost (exponent = 1.5), and the reciprocal of the square of predicted cost (exponent = 2.0). We also considered the reciprocal of the natural logarithm of predicted cost.

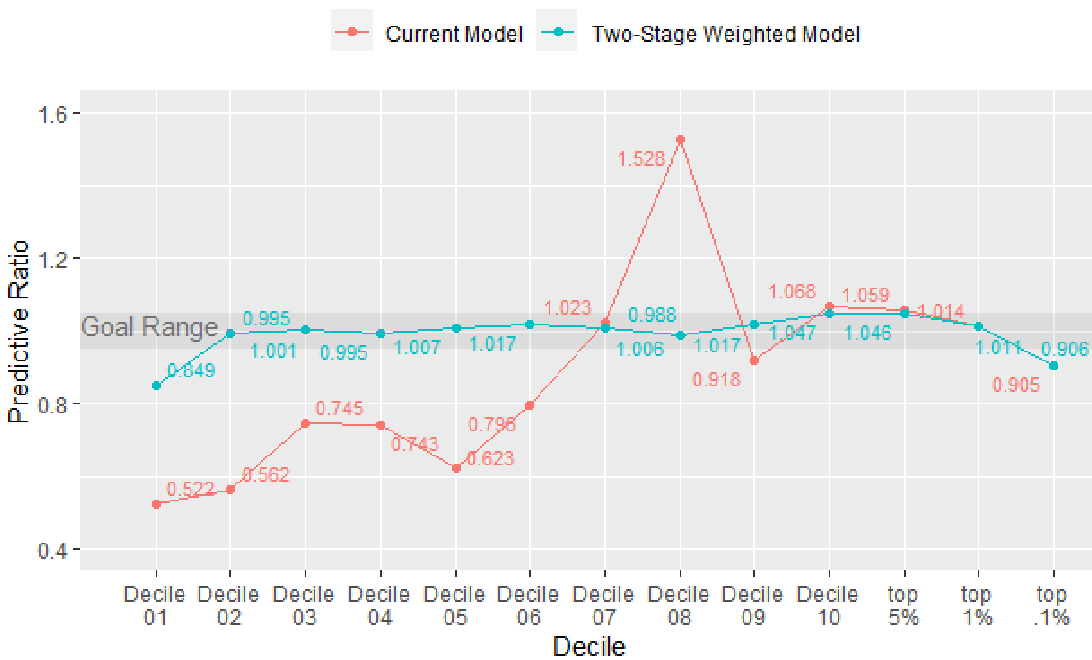
We created weighted least squares estimates of the HHS risk adjustment adult and child models using each of the series of weights, and computed PRs for each estimated model. We evaluated the models based on their PRs, particularly examining lower predicted expenditure deciles. We found that weighting factors by the reciprocal of predicted cost best achieved the objective of improving the prediction for the lowest-risk enrollee subpopulations. Further, we found that capping predicted cost at its 2.5^{th} and 97.5^{th} percentiles in the adult model and the 2.5^{th} and 99.5^{th} percentiles in the child models avoided excessively large or small weights for any observations, and improved the model fit. We multiplicatively combined this new weighting factor with the existing enrollment duration fraction weight[62] to form our final regression weight. Using this final structure for the two-stage weighted approach, we further considered the impacts of the two-stage weighted model specification.

[62] The eligibility fraction is person-specific weight where the number of months enrolled is divided by 12.

2.3 Impact of the Two-Stage Weighted Approach

As part of our consideration of the two-stage weighted approach, we considered its impact on a variety of different subpopulations and other trade-offs. The main purpose of the two-stage weighted approach is to improve prediction of the current models among lower-risk enrollees, which the two-stage weighted approach achieves. As seen in Figure 2.2, using the 2018 enrollee-level EDGE dataset, the two-stage weighted approach improves the PRs of the lower deciles compared to the current models. Similar results can be also seen when using the 2016 and 2017 EDGE data. Figure 2.2 also shows that the two-stage weighted approach eliminates the overprediction observed in risk decile 8 and slightly reduces the overprediction observed in risk decile 10.

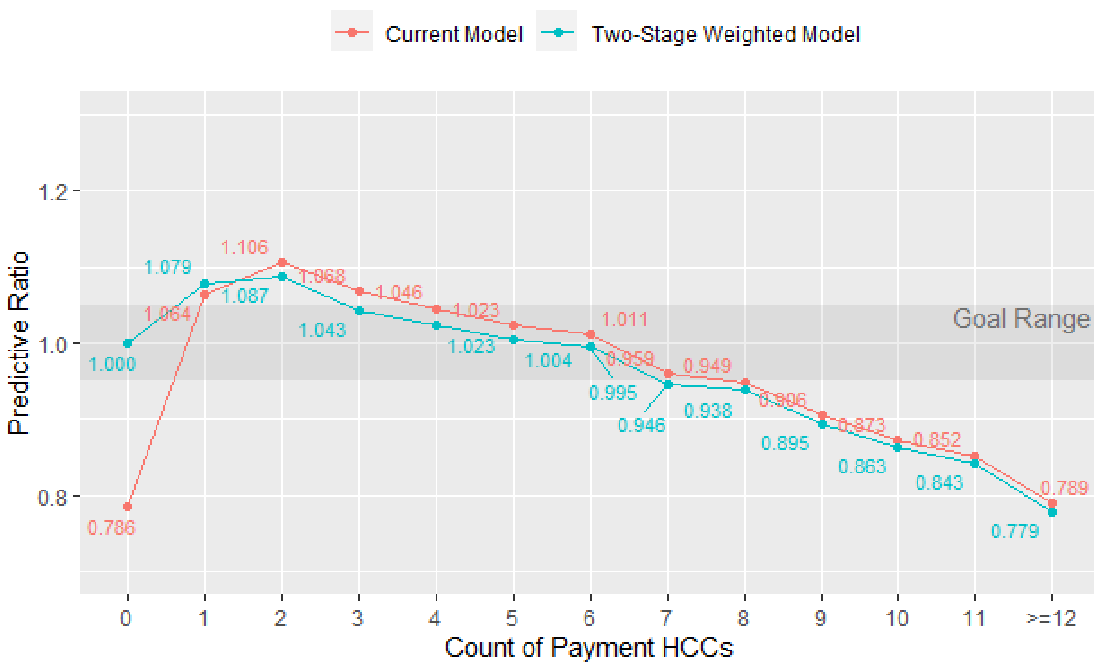
Figure 2.2 Adult Silver Plan Model Predictive Ratios by Decile, 2018 Enrollee-Level EDGE Dataset[63]



[63] In isolation from other model changes, the two-stage weighted approach overpredicts plan liability across the entire sample by 3.6 percent and, when combined with the other model specification changes outlined in Chapters 3 and 4, overpredicts the entire sample by 1.16 percent. Due to this limited impact and for simplicity in understanding the results and analyses shared in this paper, the risk scores and PRs for the two-stage weighted approach and the combined approach discussed in Chapter 5 are not normalized to be 1.00.

We also found that the two-stage weighted approach improves the PRs for enrollees without HCCs, as seen in Figure 2.3. The PRs for enrollees with 1, 2, or any other number of HCCs are only slightly affected by the two-stage weighting approach.

Figure 2.3: Adult Silver Plan Model Predictive Ratios by Number of HCCs, 2018 Enrollee-Level EDGE Dataset



After considering these gains in improved prediction for the lowest-risk enrollees (enrollees with no HCCs and low-cost enrollees), we also considered whether the two-stage weighted approach worsens the fit of the model along other dimensions. We identified three key areas where the two-stage weighted approach worsens the fit of the applicable model for certain subpopulations. First, as seen in Figure 2.4, the two-stage weighted approach predicts plan liability by age-sex factor less accurately than the current model, especially for younger and older women. Overall, we considered this to be an acceptable trade-off because across all age and sex factors, most PRs were within a tolerable threshold of +/– 5 percent (e.g., 0.95 to 1.05) and as seen Figure 2.5, the two-stage weighted approach had the major benefit of more accurately predicting the age-sex factors for the enrollees without HCCs, which is a much larger population than the enrollees with HCCs.

Figure 2.4: Adult Silver Plan Model Predictive Ratios by Age-Sex Factor, 2018 Enrollee-Level EDGE Dataset

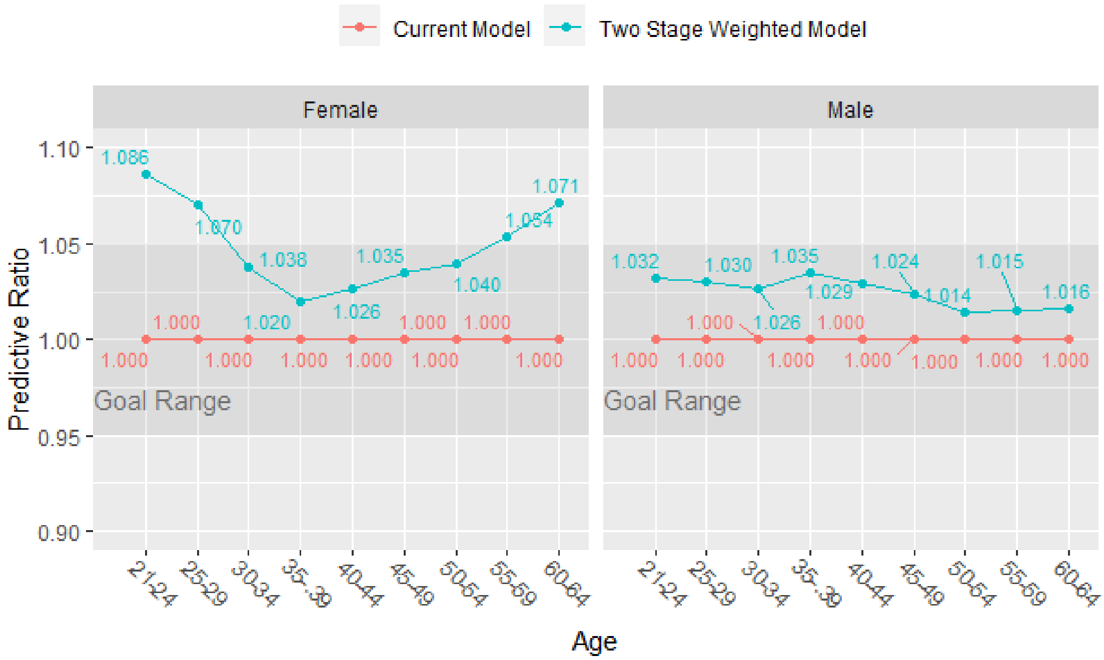
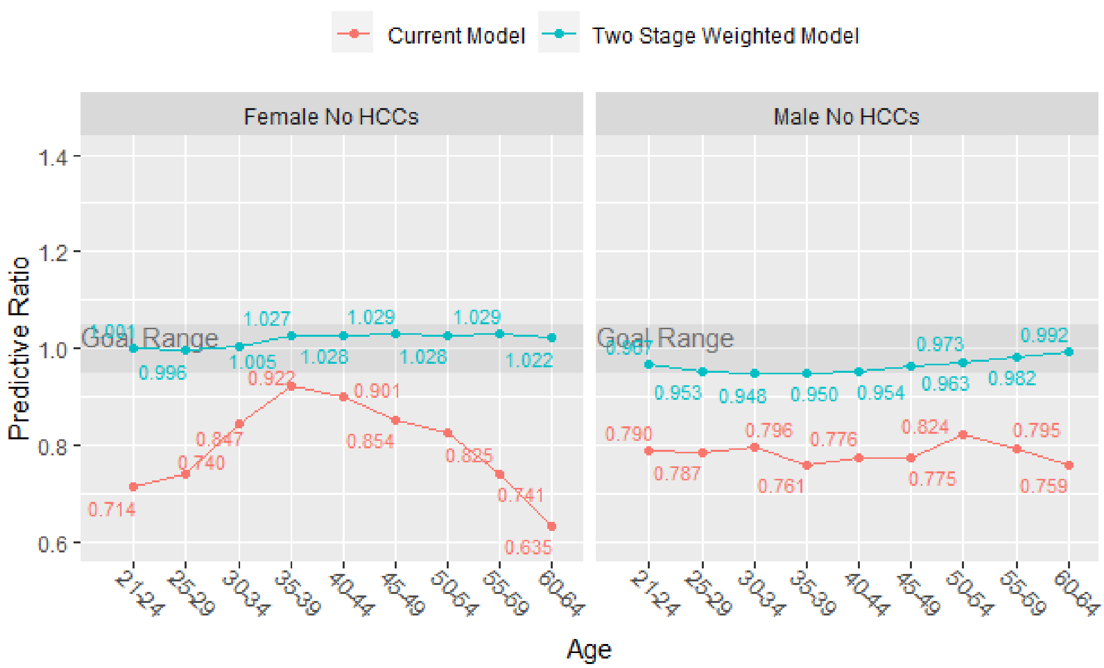


Figure 2.5: Adult Silver Plan Model Predictive Ratios by Age-Sex Factor for Enrollees without HCCs, 2018 Enrollee-Level EDGE Dataset



Second, the two-stage weighted approach is less accurate at predicting costs for certain HCCs, as seen in Table 2.1. The two-stage weighting worsened the adult model PRs by at least

5 percent in 14 (out of 91) ungrouped HCCs and 3 (out of 18) HCC groups.[64] Again, we considered this reduced accuracy to be an acceptable trade-off because most of the PRs for the two-stage weighted approach remained within a tolerable threshold of +/– 5 percent (e.g., 0.95 to 1.05), most enrollees do not have HCCs, and the two-stage weighted approach predicts plan liability better for those enrollees.

Table 2.1 HCCs and HCC Groups Whose PR in the Silver Metal Level Adult Models Worsens by at least 5 Percentage Points (pp) Under the Two-Stage Weighted Approach, 2018 Enrollee-Level EDGE Dataset

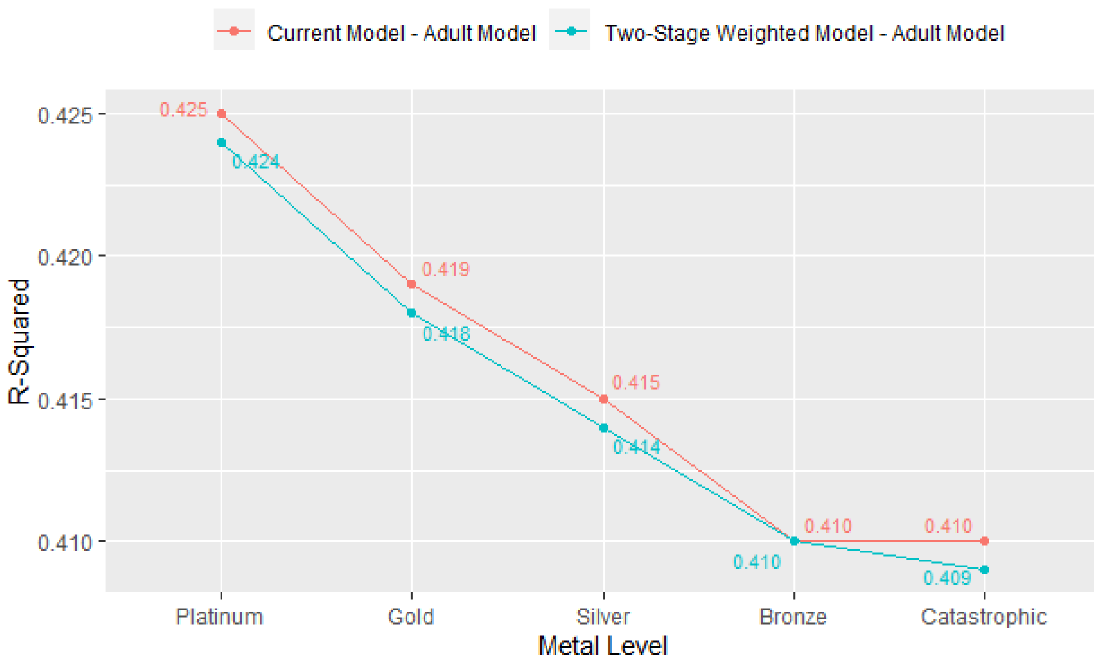
Category	*Label*	*PR*	*pp Difference*
HCC 30	Adrenal, Pituitary, and Other Significant Endocrine Disorders	0.948	5.2
HCC 63	Cleft Lip/Cleft Palate	0.944	5.6
HCC 75	Coagulation Defects and Other Specified Hematological Disorders	0.936	6.4
HCC 87_2	Delusional and Other Specified Psychotic Disorders, Unspecified Psychosis	0.943	5.7
HCC 94	Anorexia/Bulimia Nervosa	0.938	6.2
HCC 97	Down Syndrome, Fragile X, Other Chromosomal Anomalies, and Congenital Malformation Syndromes	0.907	9.3
HCC 102	Autistic Disorder	0.889	11.1
HCC 112	Quadriplegic Cerebral Palsy	1.145	14.5
HCC 142	Specified Heart Arrhythmias	0.942	5.8
HCC 151	Monoplegia, Other Paralytic Syndromes	0.949	5.1
HCC 210	(Ongoing) Pregnancy without Delivery with Major Complications	1.059	5.9
HCC 211	(Ongoing) Pregnancy without Delivery with Complications	1.074	7.4
HCC 212	(Ongoing) Pregnancy without Delivery with No or Minor Complications	1.105	10.5
HCC 219	Major Skin Burn or Condition	0.925	7.5
G07A	Sickle Cell Anemia/Beta Thalassemia	0.926	7.4
G17A	Miscarriage	1.089	8.9
G21	Congenital Heart/Circulatory Disorders	0.896	10.4

Third, the two-stage weighted approach had slightly lower R-squared values compared to the current models. As seen in Figure 2.6, there is a minor decrease in the R-squared using the two-stage weighted approach compared to the current models. The decrease in R-squared is at most 0.1 percentage points for all metals. Similar to the worsening of the age-sex cell and the HCC PRs, we were not concerned about the lower R-squared as the reduction in fit was minor at

[64] For the vast majority of HCCs, the impact is very small and most affected HCCs or HCC groups have very small sample sizes. For example, only one HCC or HCC group in Table 2.1 was present in greater than 1 percent of the enrollees in the 2018 enrollee-level EDGE dataset (HCC 142 *Specified Heart Arrhythmias*). All other HCCs in the table had recalibration dataset frequencies of less than 0.5 percent of enrollees.

all metal levels and the two-stage weighted approach better predicts plan liability for enrollees with no HCCs, which is the majority of enrollees.

Figure 2.6 Comparison of the R-squared Statistic in the Current Adult Model vs. the Two-Stage Weighted Approach, 2018 Enrollee-Level EDGE Dataset



After considering this analysis and the impact on model performance, including the impact on the models' PRs and R-squared statistics, we determined that the two-stage weighted approach does not have material unintended consequences in model performance along other dimensions.

As previously discussed, one of the key benefits of the two-stage weighted approach, which is conducted during recalibration, is that it does not require changes to the number and type of factors in the HHS risk adjustment models and does not add additional complexity to the models. We also found that the two-stage weighted approach did not meaningfully change the coefficients for most HCCs, unlike the non-linear and full HCC counts approaches. For these reasons, we found that the trade-offs in model improvement for the lowest-risk enrollee subpopulations were worth slightly worsening the model fit in other areas. As described in later chapters of this paper, we are also considering other changes to improve model fit for other subpopulations, which would offset the reduction in R-squared and many of the worsened PRs from the two-stage weighted approach.

The two-stage weighted approach effectively eliminates underprediction for the lowest-risk enrollees.[65] The two-stage weighted approach also effectively eliminated overprediction for

[65] As we discuss in Chapter 4, we are considering a separate approach to improve the models' predictive accuracy for the very highest-risk enrollees.

partial-year adult enrollees without HCCs (e.g., enrollees in risk decile 8 as seen in Figure 2.2). It does so because it turns the estimates of the current enrollment duration factors to all be (slightly) negative, reducing their overprediction. This is not a particularly desirable solution, however, because it does not address, and even exacerbates, the underprediction for partial-year adult enrollees with HCCs. For example, the underprediction for adult silver enrollees with HCCs and with 1 to 3 months of enrollment increases from 37 percent to 40 percent.[66] Chapter 3 addresses prediction issues with the current enrollment duration factors, including options to improve predictive accuracy for partial-year adult enrollees.

2.4 Conclusion

In conclusion, we have considered a variety of different options to improve the models' predictive accuracy for the lowest-risk enrollees (that is, enrollees without HCCs and low-cost enrollees), including creation of separate age-sex factors for enrollees with and without HCCs, a non-linear approach, a full HCC counts approach, and the two-stage weighted approach. We determined that the two-stage weighted approach can improve prediction for the lowest-risk enrollees with limited trade-offs in other parts of the models' performance.

[66] As we discuss in Chapter 3, we are considering a separate approach to improve the models' predictive accuracy for partial-year adult enrollees.

Chapter 3: Improving the Models' Predictive Accuracy for Partial Year Enrollees – Enrollment Duration Factors

This chapter describes the current enrollment duration factors in the adult models and discusses potential future changes to these factors, including the proposed 2022 Payment Notice updates to the enrollment duration factors. We also review our testing of other options for the enrollment duration factors.

3.1 Background on the Enrollment Duration Factors[67]

Partial-year enrollment has been common in both the individual and small group (including merged) markets since at least the establishment of the Exchanges, if not earlier. As seen in the comparison of several years of enrollee-level EDGE datasets in Table 3.1 below, for the 2016 benefit year, 50 percent of adult enrollees had fewer than 12 months of enrollment.[68] This percentage has been decreasing over time; for the 2019 benefit year, 41 percent of all adult enrollees had fewer than 12 months of enrollment.

Table 3.1: Adult Enrollment Duration, 2016–2019 Enrollee-Level EDGE Datasets

	2016 EDGE		2017 EDGE		2018 EDGE		2019 EDGE	
1 month	1,491,672	7%	1,307,825	6%	1,092,995	6%	1,054,267	6%
2 months	1,138,150	5%	1,073,182	5%	958,999	5%	875,051	5%
3 months	1,054,764	5%	990,728	5%	866,983	4%	817,502	4%
4 months	960,115	5%	907,115	4%	870,519	4%	793,369	4%
5 months	825,593	4%	787,706	4%	731,123	4%	704,162	4%
6 months	769,669	4%	726,128	4%	684,214	4%	653,851	3%
7 months	707,122	3%	658,367	3%	604,500	3%	576,350	3%
8 months	666,268	3%	626,809	3%	598,107	3%	555,080	3%
9 months	664,113	3%	586,373	3%	545,014	3%	510,879	3%
10 months	1,037,162	5%	857,849	4%	567,996	3%	491,573	3%
11 months	1,215,181	6%	1,052,204	5%	707,796	4%	628,111	3%
12 months	10,507,809	50%	11,161,412	54%	11,135,210	58%	11,059,922	59%

To develop the current enrollment duration factors for the adult models, as we discussed in the 2016 Risk Adjustment White Paper,[69] we reviewed the annualized predicted expenditures, actual expenditures, and PRs by enrollment duration groups (for each: 1 month, 2 months, and so

[67] See supra note 46 for a summary of our analysis with respect to use of enrollment duration factors in the child and infant models. Also see 85 FR at 7103 and 7104. Our current focus is similarly on the adult models and improving prediction of risk for partial-year adult enrollees.
[68] If an enrollee is enrolled in non-calendar year small group coverage, as long as they remain with the same issuer throughout the benefit year, they will be considered a full-year enrollee. This is true even if they switch plans, as long as they remain with the same issuer. However, if they switch to a plan with a different issuer, they will be considered a partial-year enrollee due to data limitations that prevent HHS from knowing that they are the same enrollee with full coverage throughout the benefit year who just switched issuers between plan years.
[69] See pages 35-39 of the March 2016 Risk Adjustment Methodology White Paper (March 24, 2016), available at: https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf.

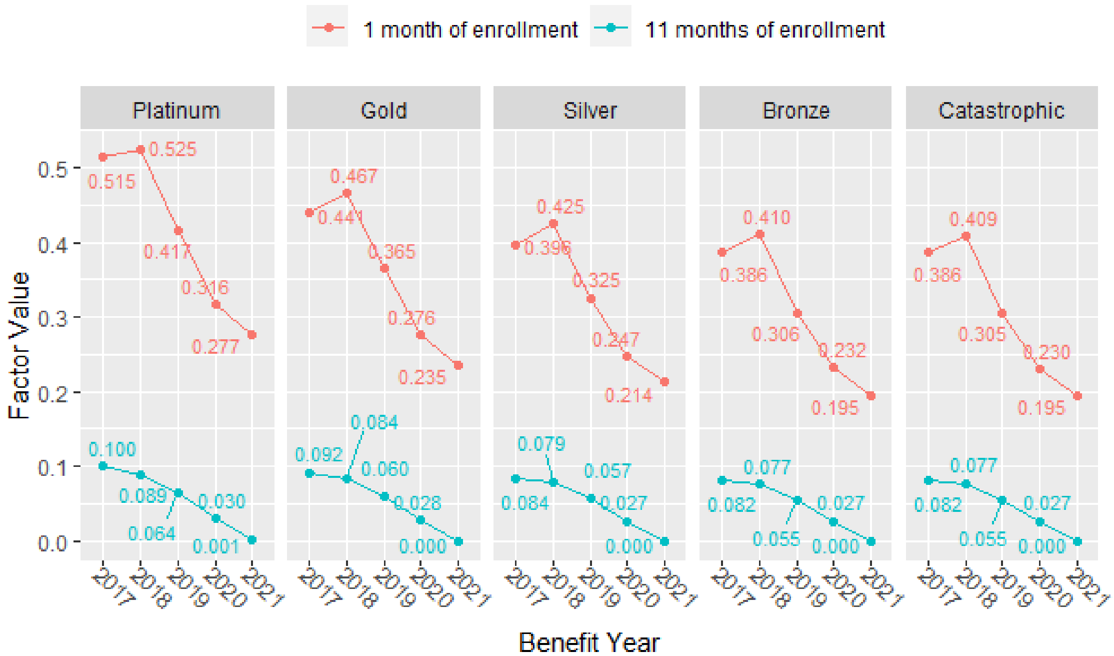
on up to 12 months) for our risk adjustment concurrent modeling sample, which was made up of adults in the 2014 MarketScan® data. We found that actuarial risk for adult enrollees with short enrollment periods tended to be underpredicted in our methodology, and actuarial risk for adult enrollees with full enrollment periods (12 months) tended to be overpredicted.

Many commenters to the 2016 White Paper expressed a preference for adding enrollment duration binary indicator variables (indicating enrollment duration of: 1 month, 2 months, and so on up to 11 months[70]) as additional risk factors, as opposed to creating separate models based on enrollment duration. After reviewing this feedback, we proposed and finalized in the 2018 Payment Notice that, beginning for the 2017 benefit year, the adult models would include enrollment duration factors that apply to all adults with partial-year enrollment.[71] Although the value for the factors change from year to year due to the annual model recalibration, we have not made changes to the structure of the enrollment duration factors since they were first adopted.

The values for the enrollment duration factors have generally decreased since they were first introduced to the HHS-operated risk adjustment adult models for the 2017 benefit year, reflecting a reduced impact of enrollment duration on risk scores of partial year enrollees. After a slight increase between 2017 and 2018, the factors have decreased significantly from 2018 to 2021, and in some cases (the 10- and 11-month factors) the factors are now 0.000, relative to a 12-month enrollment baseline.[72] Figures 3.1 and 3.2 below show the change in the values of the 1-month and 11-month enrollment duration factors over time. The values of the 2- through 10-month enrollment duration factors show a similar decrease over the same timeframe.

[70] Twelve months is the reference group and therefore is not included.
[71] 81 FR 94058 at 94071 – 94074.
[72] In unconstrained models, these factors are negative; therefore, we constrained them to zero because we do not believe negative enrollment duration factors are appropriate, as this would create inappropriate incentives.

Figure 3.1: 1- and 11-Month Enrollment Duration Factors, 2017–2021



The decline in the value of the enrollment duration factors occurred while HHS transitioned from using MarketScan® data to using enrollee-level EDGE data for the annual recalibration of the risk adjustment models, a transition which began for the 2019 and 2020 benefit year recalibrations, and was completed for the 2021 benefit year recalibration. The availability of enrollee-level EDGE data enabled HHS to further analyze this decline and begin to consider potential changes to how we account for partial-year enrollment in the adult risk adjustment models.

As described in section 1.4, the current enrollment duration factors, which are estimated on all partial-year enrollees (with and without HCCs), have factors fairly close to zero. The largest value for the enrollment duration factors in the 2021 benefit year models are around 0.200. These coefficients reflect a weak relationship between enrollment duration and costs when estimating the relationship using enrollees with and without HCCs and controlling for all other variables in the 2021 models. The weak relationship masks the different relationships between enrollment duration and costs that exist for partial-year enrollees with HCCs and partial-year enrollees without HCCs. In particular, there is a strong negative relationship between enrollment duration and costs[73] for enrollees with HCCs which is offset by a weak positive relationship between enrollment duration and costs for enrollees without HCCs.

While we are not certain why enrollment duration is only inversely related to monthly costs among adults with HCCs, the most plausible hypothesis is that medical treatment for many

[73] In the risk adjustment models, costs (or expenditures) are annualized, meaning that an enrollee who incurs \$5,000 of costs in 5 months would have annualized costs of (12/5)*\$5,000 = \$12,000. The negative relationship between enrollment duration and costs discussed in this section is predicated on this annualization.

HCC diagnoses likely has a "fixed cost" element that does not vary with the number of months of enrollment. This fixed cost gets spread over a larger number of months as enrollment duration increases, decreasing average monthly expenditures. This is particularly true for acute or time-limited HCCs (e.g., HCC 212 (Ongoing) Pregnancy without Delivery with No or Minor Complications), including HCCs implying definitive surgical treatment. For example, a 1-month enrollee with an HCC likely had active treatment for, or influenced by, that HCC in their 1 month of enrollment. Conversely, a 12-month enrollee with an HCC may only have had active treatment for, or influenced by, that HCC in 1 of their 12 enrollment months. These two enrollees might have the same total costs for active treatment of the HCC, but the per-month costs for the 1-month enrollee would be 12 times higher than the per-month costs for the 12-month enrollee. However, the cause remains uncertain as this explanation is unlikely to fully account for the inverse relationship between enrollment duration and monthly costs among adults with HCCs.

As described in section 1.4, we have been exploring ways to improve prediction of costs for partial-year adult enrollees with HCCs and partial-year adult enrollees without HCCs under the current enrollment duration factors. We describe some potential changes to the enrollment duration factors to improve the prediction of partial-year enrollees in the following section.

3.2 Potential Changes to the Enrollment Duration Factors

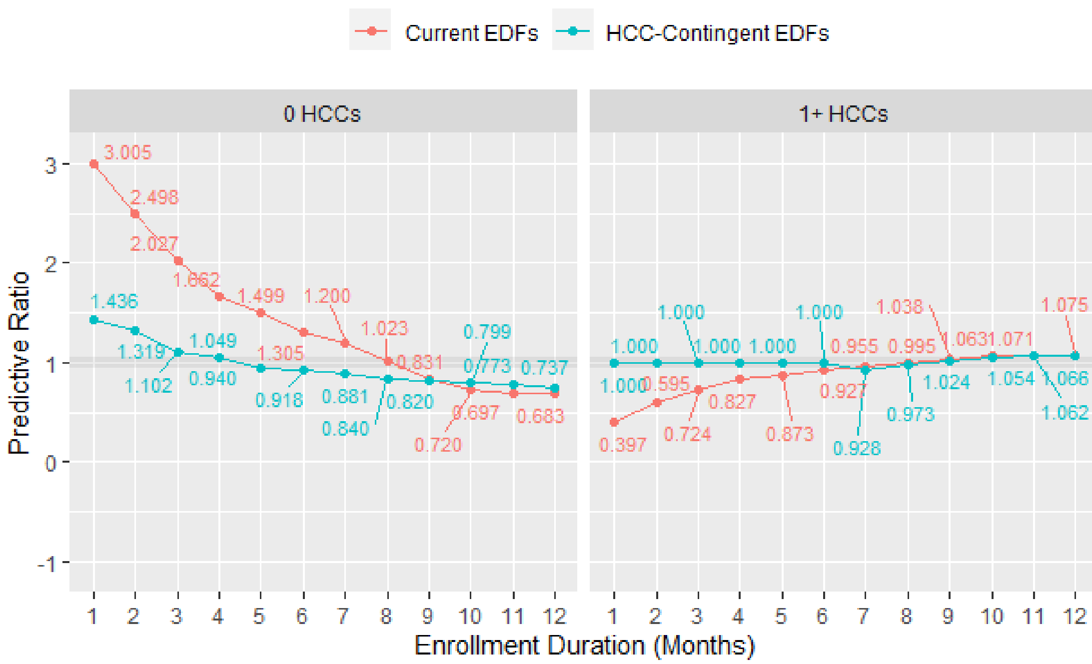
HHS considered a variety of options to improve the prediction of partial-year adult enrollees through modifying the enrollment duration factors. We found that enrollment duration factors that were contingent on whether an enrollee had one or more payment HCCs were the best predictor of a meaningful cost distinction and best reflected the relationship between enrollment duration, HCCs, and cost. The HCC-contingent enrollment duration factors that were proposed in the 2022 Payment Notice,[74] but not finalized, are six binary variables that flag enrollment durations up to six months only among enrollees with at least one HCC.

In this chapter, we report on analysis of the estimated impact of the HCC-contingent enrollment duration factors that were proposed in the proposed 2022 Payment Notice, in isolation (that is, without the two-stage weighting approach described in Chapter 2 or the interacted HCC counts model specification described in Chapter 4). The PRs associated with the HCC-contingent enrollment duration factors show a significant improvement over the current factors' PRs for partial-year adult enrollees with and without HCCs, as shown in Figure 3.2 below. The HCC-contingent enrollment duration factors address the underprediction of plan liability for adults with short enrollment duration and at least one payment HCC, as well as the overprediction of plan liability for adults with short enrollment duration and no payment HCCs. Using the 2018 enrollee-level EDGE dataset to analyze the adult silver model, the HCC-contingent enrollment duration factors eliminate the underprediction among adults with HCCs and 1-6 months of enrollment.[75] Using the HCC-contingent enrollment duration factors, the

[74] 85 FR 78572 at 78585 – 78586.
[75] When the HCC-contingent enrollment duration factors are combined with the two-stage weighted model specification (see Chapter 5), these underpredictions are reduced rather than eliminated.

overprediction among adult silver plan enrollees without HCCs with 1-4 months of enrollment duration is substantially reduced on average and eliminated by 5 months of enrollment duration.

Figure 3.2: Adult Silver Plan Model Predictive Ratios by Enrollment Length Using Current and HCC-Contingent Enrollment Duration Factors, 2018 Enrollee-Level EDGE Dataset



The HCC-contingent enrollment duration factors stop at six months, because the monthly average cost variation by number of months enrolled is substantially reduced after six months, which we observed consistently across data years. The underprediction among adults without HCCs with 7-12 months of enrollment duration is consistent with the broader underprediction of full year enrollees without HCCs discussed in Chapter 2.[76] We also tested adding new enrollment duration factors up to nine months, but did not observe any significant changes when compared to the proposed factors for 1-6 months, and the factors for the 7-9 months were close to zero and sometimes negative. We also did not observe any material improvement to PRs by adding 7-9 month enrollment duration factors. For these reasons, we determined the best option to improve prediction for partial-year adult enrollees is by using the HCC-contingent enrollment duration factors of 1-6 months as proposed in the proposed 2022 Payment Notice.

3.3 Testing Other Options for the Enrollment Duration Factors

Based on suggestions from commenters to past rules and our own continuing analysis, we also tested several other options to restructure the current enrollment duration factors. We considered varying the enrollment duration factors by type of partial-year enrollment (enrolling after open enrollment through a special enrollment period versus enrolling during open

[76] The two-stage weighted approach, discussed in detail in sections 2.2 and 2.3, helps mitigate the underprediction of full year enrollees without HCCs.

enrollment and dropping enrollment partway through the year); by market (individual versus small group market); and by specific HCC (as well as by type of HCC – acute versus chronic). While we are not considering these types of changes to the enrollment duration factors at this time, our analyses of these options are described in this section.

3.3.1 Partial-Year Enrollment Type

Some commenters have expressed concerns that special enrollment period (SEP) enrollees, especially in the individual market, are underpredicted in the current risk adjustment models.[77] In the 2021 Payment Notice,[78] we explained that our preliminary analysis of the 2017 enrollee-level EDGE dataset found that, in comparison to the effect of the presence of HCCs on enrollment duration factors, enrollment timing (e.g., enrollment at the beginning of the year compared to enrollment after open enrollment period, or drop in enrollment before the end of the year) did not appear to affect PMPM expenditures on average. Nonetheless, we tested adding variables to the model for SEP enrollment (first enrollment after January of the applicable benefit year) and drop in enrollment (last enrollment before December of the applicable benefit year) by market.[79] Under this simulation, these variables were added to an adult enrollee's enrollment duration factor as an additional component of their risk score. We repeated this analysis using the 2019 enrollee-level EDGE dataset and found similar results. The estimates for these variables using the 2017 and 2019 enrollee-level EDGE datasets are shown in Table 3.2 below.

Table 3.2: Partial-Year Enrollment Type Factors

	2017 EDGE Data		2019 EDGE Data	
Label	**Individual Market**	**Small Group Market**	**Individual Market**	**Small Group Market**
SEP	0.0685	-0.1099	0.0643	-0.0810
DROP	-0.0484	0.0902	-0.0452	0.0667

Note: SEP is an additional effect for enrollees with first enrollment after March (for 2017 EDGE data) or January (for 2019 EDGE data);[80] DROP is an additional effect for enrollees with last enrollment before December. The SEP and DROP variables were estimated in separate models.

As seen in the table, the additive factors for partial-year enrollment type would be very close to zero and in some cases negative. The additional findings from the analysis of the 2019 enrollee-level EDGE dataset were consistent with our prior findings that enrollment timing does

---

[77] Some researchers have also found that special enrollment period enrollees have higher costs than open enrollment enrollees. See Garabedian et al., Costs Are Higher For Marketplace Members Who Enroll During Special Enrollment Periods Compared With Open Enrollment, Health Affairs, August 2020, available at: https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2019.01155.

[78] 85 FR 29164 at 29189.

[79] For purposes of this testing, we used the current enrollment duration factors.

[80] Open enrollment for the 2017 benefit year extended through January 31, 2017, so our analysis of the 2017 enrollee-level EDGE dataset includes enrollees with first enrollment after March in the SEP variable. Open enrollment for the 2019 benefit year extended through December 15, 2018, so our analysis of the 2019 enrollee-level EDGE dataset includes enrollees with first enrollment after January in the SEP variable.

not appear to affect PMPM expenditures on average. Therefore, we have not pursued proposals to vary the enrollment duration factors by partial-year enrollment type.

3.3.2 Market

HHS has also received comments requesting consideration of market-specific enrollment duration factors to account for perceived variance in risk of partial-year enrollees by market.[81] We agreed that this was an area that should be explored further, as there may be differences in partial-year enrollment patterns between the individual and small group (including merged) markets. Our preliminary analysis of the 2017 enrollee-level EDGE dataset found that separate enrollment duration factors by market in the adult models may be warranted, given the differences in risk profiles of partial-year enrollees in the individual and small group markets.[82] However, we were unable to develop and propose separate enrollment duration factors by market at that time due to limitations in the available enrollee-level EDGE data.[83] Nonetheless, we have continued to consider these issues.

As we stated in the proposed 2022 Payment Notice,[84] as part of our updated analysis on partial-year enrollment using the 2018 enrollee-level EDGE dataset, we considered adoption of enrollment duration factors by market. However, we did not find a meaningful distinction in relative costs between markets on average once we implemented the proposed HCC-contingent enrollment duration factors in the proposed 2022 Payment Notice. Therefore, we determined it would not be necessary to introduce market-specific factors if HCC-contingent enrollment duration factors were implemented. Even though reasons for and patterns of partial-year enrollment differ by market, we concluded that the patterns most relevant for predicting cost (e.g., how enrollment duration relates to cost conditional on the presence of HCCs) were the same for both markets. Table 3.3 shows potential HCC-contingent enrollment duration factors by market for adult silver plan enrollees using the 2017, 2018 and 2019 enrollee-level EDGE datasets. As seen in Table 3.3, if the HCC-contingent factors were to vary by market, the factors for both markets would generally be very similar, which would add little value to the models while adding additional complexity.

[81] See 86 FR 24140 at 24162 and 85 FR 29164 at 29190.
[82] In the enrollee-level EDGE dataset, merged market enrollees are assigned to the individual or small group market indicator based on their plan.
[83] See 85 FR at 7103-7104.
[84] 85 FR 78572 at 78585.

Table 3.3: Enrollment Duration Factors for Adult Silver Plan Enrollees by Market

	2017 EDGE		2018 EDGE		2019 EDGE	
Label	**Individual Market**	**Small Group Market**	**Individual Market**	**Small Group Market**	**Individual Market**	**Small Group Market**
Enrolled for 1 month, at least one payment HCC	6.766	6.453	6.396	7.823	6.026	6.134
Enrolled for 2 months, at least one payment HCC	2.066	2.794	1.853	2.608	1.755	2.053
Enrolled for 3 months, at least one payment HCC	1.213	1.331	0.874	1.280	1.953	1.385
Enrolled for 4 months, at least one payment HCC	0.510	0.747	0.303	0.615	0.511	0.764
Enrolled for 5 months, at least one payment HCC	0.244	0.419	0.186	0.505	1.315	0.330
Enrolled for 6 months, at least one payment HCC	0.132	0.335	-0.063	0.336	-0.061	0.345

3.3.3 Specific HCCs

HHS also received comments suggesting that certain HCCs may be more common among partial-year enrollees than others and may be contributing more heavily to partial-year enrollee costs. As part of our consideration of these comments, we analyzed the prevalence of HCCs among enrollees with one month of enrollment duration.[85] Some commenters asked HHS to consider whether enrollment duration factors should be tied to certain HCCs, believing that not all HCCs contribute equally to the coefficient for enrollees with the 1-month enrollment duration factor. These stakeholders requested that we constrain the enrollment duration factor to a subset of HCCs driving the high 1-month enrollment duration factor value in the proposed 2022 Payment Notice. One commenter recommended HCC-specific enrollment duration factors for maternity HCCs be finalized for the 2022 benefit year.

As part of our analysis of the proposed HCC-contingent enrollment duration factors in the proposed 2022 Payment Notice, we reviewed the most common HCCs in the 2018 enrollee-level EDGE dataset for 1-month enrollees to determine whether the HCC profile of partial-year enrollees with these common HCCs might differ substantially from full-year enrollees (that is, whether partial-year enrollees tend to have different HCCs than full-year enrollees). As seen in columns 1 and 2 of Table 3.4, our analysis showed that the HCCs that tend to be the most common among 1-month adult enrollees also tend to be the most common HCCs among all adult enrollees with a few exceptions. Additionally, as seen in column 3 of Table 3.4, we also assessed whether there were HCCs that tended to be more common among 1-month adult enrollees compared to all adults and identified those HCCs. As shown in Table 3.4 below, the only

[85] We focused the analysis on enrollees with one month of enrollment because some commenters to the proposed 2022 Payment Notice noted the large proposed 1-month enrollment duration factor under the proposed new HCC-contingent enrollment duration factors and expressed concerns that not all HCCs would equally contribute to that factor.

condition that is both most common among 1-month adult enrollees and more common among 1-month enrollees as compared to all adult enrollees, is HCC 212 (Ongoing) Pregnancy without Delivery with No or Minor Complications. This implies that HCC profiles are more similar than different between partial-year enrollees and full-year enrollees.

Table 3.4: Common HCCs Among 1-Month Enrollees, 2018 Enrollee-Level EDGE Dataset

Most common among 1-month adult enrollees	Most common among all adult enrollees	More common among 1-month adult enrollees compared to all adult enrollees
HCC 20 Diabetes with Chronic Complications	HCC 20 Diabetes with Chronic Complications	HCC 9 Lung, Brain, and Other Severe Cancers, Including Pediatric Acute Lymphoid Leukemia
HCC 21 Diabetes without Complication	HCC 21 Diabetes without Complications	HCC 127 Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes
HCC 88 Major Depressive Disorder, Severe, and Bipolar Disorders	HCC 88 Major Depressive Disorder, Severe, and Bipolar Disorders	HCC 184 End-Stage Renal Disease
HCC 161_2 Asthma, Except Severe	HCC 160 Chronic Obstructive Pulmonary Disease, Including Bronchiectasis	HCC 211 (Ongoing) Pregnancy without Delivery with Complications
HCC 212 (Ongoing) Pregnancy without Delivery with No or Minor Complications	HCC 161_2 Asthma, Except Severe	HCC 212 (Ongoing) Pregnancy without Delivery with No or Minor Complications

We also categorized HCCs into acute versus chronic HCCs and tested how enrollment duration patterns varied for those two groups, as we understood commenters' concern that the pattern of PMPM costs by enrollment duration could vary by HCC. As part of this analysis, we considered whether adults with any acute HCCs included a significantly higher fraction of enrollees with short enrollment duration or had significantly higher costs than adults with only chronic HCCs. As seen in Table 3.5, in the 2017 enrollee-level EDGE dataset, the distribution of enrollment duration and PMPM allowed charges by enrollment duration is similar for adults with any acute HCCs versus adults with only chronic HCCs. In other words, the general pattern of decreasing PMPM costs with higher enrollment duration held true, regardless of whether HCCs were acute or chronic.

Table 3.5: PMPM Allowed Charges by Enrollment Duration and Presence of Acute versus Chronic HCCs, 2017 Enrollee-Level EDGE Dataset

	Adults with Any Acute HCCs		Adults with Chronic HCCs Only	
Enrollment Duration	**Number of Enrollees**	**PMPM Allowed Charges**	**Number of Enrollees**	**PMPM Allowed Charges**
1 month	7,567	$16,402	41,449	$16,474
2 months	11,448	$10,604	62,141	$11,268
3 months	16,044	$8,343	81,809	$8,815
4 months	17,918	$6,673	91,817	$7,807
5 months	19,360	$6,014	91,997	$6,688
6 months	21,154	$5,116	94,907	$6,164
7 months	21,593	$4,635	92,226	$5,468
8 months	22,917	$4,162	93,351	$4,991
9 months	23,663	$3,995	94,012	$4,698
10 months	34,520	$3,599	144,179	$4,761
11 months	45,505	$3,332	195,130	$4,781
12 months	468,931	$3,087	2,252,102	$5,706

Note: The list of HCCs we considered to be acute for the purposes of this analysis may be found in Appendix B.

We, therefore, determined that it would add unnecessary complexity to the adult models to introduce one set of enrollment duration factors for adults with any acute HCCs and a separate set for adults with only chronic HCCs. As explained in the 2022 Payment Notice,[86] our primary concern with tying the enrollment duration factors to specific HCCs, such as the maternity HCCs, is that several new factors would have to be added to the models to create these HCC-specific enrollment duration factors, adding an additional level of complexity and potential instability to the model that may not be offset by any gains in predictive accuracy. We continue to believe that the most important factor in predicting costs for partial-year enrollees is the presence of any HCC, rather than the presence of specific types of HCCs.

3.4 Obtaining Diagnosis Codes for Partial-Year Enrollees

Some commenters to the proposed 2022 Payment Notice expressed concerns that issuers, particularly small group market issuers, small issuers, or Medicaid issuers, may have partial-year enrollees with HCCs that are not coded. These commenters expressed concerns that these issuers may have difficulty obtaining diagnoses for partial-year enrollees, creating cases where the issuer may pay claims, and incur costs, for services associated with a condition for the partial-year enrollee, but the issuer's limited time with the partial-year enrollee may not be adequate to capture the diagnosis code associated with the HCC.

[86] 86 FR 24140 at 24162.

Note that this issue differs from the case where issuers may not have a complete diagnostic profile for a partial year enrollee because the services received were not related to the diagnoses that were not captured. For example, if an enrollee received services due to a condition while enrolled with a different issuer, then the current issuer may not have all diagnosis codes for a partial year enrollee. However, such cases do not have cost implications for the current issuer since the partial year enrollee received no services associated with that diagnosis.

While there are likely some cases where an enrollee only receives risk adjustment ineligible services, our analysis found no evidence that it is associated with underpayment. As noted earlier in this chapter, the current models overpredicted for partial-year enrollees with no HCCs. Consistent with this finding, adding enrollment duration factors for partial-year enrollees without HCCs results in negative coefficients for these factors, as seen in Table 3.6. In other words, on average, costs are sufficiently low for partial-year enrollees with no HCCs that even a risk score based only on demographic factors would overpredict plan liability.

Table 3.6: Enrollment Duration Factors for Adult Silver Plan Enrollees without HCCs by Market

Factor	2017 EDGE Data		2018 EDGE Data		2019 EDGE Data	
	IND	SG	IND	SG	IND	SG
Enrolled for 1 month, zero payment HCCs	-0.079	-0.015	-0.073	-0.028	-0.054	-0.046
Enrolled for 2 months, zero payment HCCs	-0.060	-0.018	-0.058	-0.028	-0.036	-0.051
Enrolled for 3 months, zero payment HCCs	-0.037	-0.037	-0.033	-0.036	-0.024	-0.042
Enrolled for 4 months, zero payment HCCs	-0.035	-0.021	-0.025	-0.036	-0.029	-0.034
Enrolled for 5 months, zero payment HCCs	-0.036	-0.017	-0.024	-0.025	-0.025	-0.029
Enrolled for 6 months, zero payment HCCs	-0.028	-0.007	-0.030	-0.016	-0.024	-0.029

Note: IND denotes the individual market; SG denotes the small group market.

The pattern of negative factors for enrollees with short enrollment duration (1-6 months) and no HCCs is consistent with the descriptive statistics in Table 1.1 in section 1.4, which showed that short-term (1-6 month) enrollees without HCCs have lower PMPM allowed charges than short-term enrollees with HCCs or full-year enrollees without HCCs. Further, the similarity of results by market in Table 3.6 is consistent with the similarity of results by market in Table 3.3. In summary, these results support our determination that we have not found any evidence that issuers are unable to capture cost-meaningful HCCs for partial-year enrollees in the individual or small group (including merged) market.

3.5 Conclusion

In conclusion, our analyses and evaluation of the current enrollment duration factors in the adult models identified areas for improving the prediction of costs for partial-year adult enrollees with HCCs and partial-year adult enrollees without HCCs. HHS considered several possible changes to the enrollment duration factors, including varying them by partial-year enrollment type, by market, and by specific HCC or type of HCC. However, we found that the most effective change would be to make the enrollment duration factors contingent on the presence of one or more payment HCCs. HCC-contingent enrollment duration factors described in this chapter would improve the current model's predictive accuracy for these enrollee subpopulations and reduce the number of factors in the adult risk adjustment models.[87]

[87] As detailed in the proposed 2022 Payment Notice, these new factors would only apply to partial-year enrollees with up to 6 months of enrollment and at least one payment HCC. See 85 FR 78572 at 78586.

Chapter 4: Improving the Predictive Accuracy for the Very Highest-Risk Enrollees – Interacted HCC Counts

This chapter explores the prediction of risk for the very highest-risk enrollees in the current adult and child risk adjustment models. While the number of very highest-risk enrollees is small in both the adult and child models, the associated plan costs can be high and make up a significant amount of overall expenditures. In this chapter, we discuss ways to improve prediction for these enrollees through the "interacted HCC counts approach," which was discussed in the proposed 2022 Payment Notice. In addition, we provide analysis to show how the interacted HCC counts approach works and how it can improve prediction for the very highest-risk enrollees in the adult and child models.

4.1 Current Models' Prediction of Costs for the Very Highest-Risk Enrollees

As described in Chapter 1.4, the current risk adjustment models underpredicted costs for enrollees with the highest HCC counts and enrollees in the top 0.1 percent decile. This underprediction matters due to the impact the very highest-risk enrollees have on overall plan liability. While the vast majority (around 80 percent) of enrollees do not have HCCs, those with HCCs make up a significant amount of overall expenditures. For example, as seen in Tables 4.1 and 4.2 below, enrollees in the highest two risk deciles account for approximately 83 percent of total adult silver plan liability, while the top 1 percent of enrollees by predicted risk account for more than 31 percent of adult silver plan liability. While enrollees with 7 or more HCCs account for only 0.2 percent of adult enrollees, they make up more than 8 percent of adult silver plan liability.

Table 4.1: Percent of Total Adult Silver Plan Liability by Decile/Percentile, 2018 Enrollee-Level EDGE Dataset

Decile/Percentile	% Total Adult Silver Plan Liability	Cumulative % Adult Silver Plan Liability
1 (lowest)	1.36%	1.36%
2	1.95%	3.31%
3	2.11%	5.42%
4	2.62%	8.04%
5	3.99%	12.03%
6	2.85%	14.88%
7	1.40%	16.28%
8	0.82%	17.09%
9	8.62%	25.71%
10 (highest)	74.29%	100.0%
Top 5%	62.21%	-
Top 1%	31.36%	-
Top 0.1%	7.39%	-

Table 4.2: Percent of Total Adult Silver Plan Liability by HCC Count, 2018 Enrollee-Level EDGE Dataset

Number of Payment HCCs	% Total Adult Silver Plan Liability	Cumulative % Adult Silver Plan Liability
0	19.17%	19.17%
1	30.70%	49.87%
2	17.85%	67.72%
3	10.28%	78.00%
4	6.32%	84.32%
5	4.21%	88.53%
6	2.98%	91.51%
7	2.21%	93.72%
8	1.64%	95.36%
9	1.31%	96.67%
10+	3.33%	100.00%

As discussed in Chapter 1, the most expensive enrollees tend to have severe acute illness HCCs. In addition, the highest-cost enrollees are also characterized by high numbers (counts) of HCCs. As shown in Table 4.3, for enrollees with a high HCC count, there is an increasing, non-linear effect that leads to higher costs than are predicted by adding up the incremental effects of each HCC. The average PMPM actual cost for enrollees with 1 HCC is 8.4 times the average actual cost for enrollees with 0 HCCs. This ratio increases to 142.5 times for enrollees with 7 HCCs and to 313 times for enrollees with 10 or more HCCs.

Table 4.3: Ratio of Actual Plan Liability compared with Plan Liability at Zero HCCs

Number of HCCs	Ratio: Actual PL / Actual PL at 0 HCCs
1	8.355
2	18.341
3	34.354
4	55.309
5	80.489
6	108.078
7	142.520
8	171.668
9	213.788
10+	313.122

4.2 Improving Prediction for the Very Highest Risk Enrollees – Interacted HCC Counts Approach

To account for the very highest-risk enrollees, the current HHS risk adjustment adult models incorporate a severe illness adjustment that accounts for the interactions or combinations of selected HCCs. When a severity indicator is present,[88] the adult models will look for interactions with a few other HCCs or HCC groups and apply a binary indicator variable in the regression model that indicates whether the enrollee has at least one disease interaction in the category. If a disease interaction category is present, the adult enrollee receives the factors associated with the HCCs that the enrollee has, plus a severity indicator adjustment that would impact their predicted cost.[89] In the current risk adjustment adult models, the total count of an

[88] Prior to the 2021 benefit year, the HHS risk adjustment adult models included two types of severe illness HCC interactions that estimated the additional incremental costs of combinations of high severity HCCs, medium and high. In the V06 HHS-HCC classification update, the medium interaction was removed, as its parameter estimate is usually very low. For example, in the 2017 enrollee-level EDGE dataset, the medium cost severe illness interaction term's estimate was negative and constrained to $0. See Centers for Medicare & Medicaid Services, Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program (June 17, 2019), available at: https://www.hhs.gov/guidance/sites/default/files/hhs-guidance-documents/Potential%20HHS-HCC%20Updates%20for%20Risk%20Adjustment%20Program.pdf. The medium severity indicator was removed from the HHS risk adjustment adult models in the 2021 Payment Notice, see 85 FR 29164.

[89] For example, if an enrollee has HCC 2 (Sepsis), which is one of the eight severity indicators in the current adult models, which is paired with HCC 6 (opportunistic infections), which is one of the selected HCC interactions that

enrollee's HCCs does not independently affect their risk score. While this severity adjustment helps the current adult models predict costs accurately among all enrollees with qualifying severe illnesses, it still does not fully address the underprediction for the very highest-risk enrollees.

As discussed in the proposed 2022 Payment Notice,[90] we investigated alternative approaches to improve the current risk adjustment models by incorporating an HCC counts model which we believed could improve prediction for the very highest-risk enrollees while minimizing gaming concerns and disruption of current model factors. The goals for this approach were to:

1. Address the non-linearity in costs between enrollees without HCCs or with very low costs and enrollees with multiple HCCs or with high costs;
2. Empirically incorporate the cost impact of multiple complex diseases; and
3. Reduce incentives for coding proliferation to mitigate the gaming concerns with HCC counts models.

We tested different types of indicators interacted with HCC counts with the goal of improving prediction for the very highest-risk enrollees and multiple HCCs. For this approach, we assessed the HCCs for enrollees with extremely high costs. We found that many of these enrollees had HCCs among the severe illness indicators, transplant HCCs, and other HCCs related to severity of disease. In addition to these indicators, we also found that the very highest-risk enrollees were characterized by high counts of HCCs. This suggests that the count of HCCs as a measure of total disease burden *combined* with the presence of one or more severe illness HCCs holds promise as a means to improve prediction for the very highest-risk enrollees. While the current adult models' severity illness factors are intended to account for some of these severe illness interactions, our analysis found that it still underpredicts costs for the very highest-risk enrollees. As a result, in the proposed 2022 Payment Notice we proposed dropping the current severity illness indicators in the adult models and replacing them with severity and transplant indicators interacted with HCC counts factors in the adult and child models.[91]

Based on our empirical investigation of acute disease severity, including consultation with clinicians, we added 19 HCCs to our list of severe illness HCCs, and removed one of the original eight, for an expanded total of 26 severe illness HCCs, as seen in Appendix C, under the

---

would trigger a severity adjustment, the enrollee would receive the risk scores for HCC 2 and HCC 6 plus an incremental severity adjustment based on the interaction between a severe illness HCC (e.g. Sepsis) and HCC 6.

[90] 85 FR 78572.

[91] While the current risk adjustment models only apply the severity illness indicators to the adult models because the interaction terms were empirically unimportant in the child models, many of the prediction issues that we observed in the adult models were also observed in the child models. Additionally, we found that the interacted HCC counts approach improved prediction in both the adult and child models. As a result, we proposed in the proposed 2022 Payment Notice to apply the interacted HCC counts specification to both the child and adult models. We did not, however, propose to change the infant model severity illness indicators because the infant models do not categorize infant enrollees the same way as in the adult and child models, rather infant risk scores are determined by inclusion in one of 25 mutually exclusive groups, based on the infant's maturity and the severity of diagnoses, as a result a counts approach would not be applicable.

interacted HCC counts model specification proposal.[92] We experimented with interacting HCC counts separately for each of the severe illness HCCs, but we found this approach to be too complex and prone to overfitting due to estimation of a large number of parameters. We also found that we could simplify the models by interacting any of the selected severe illness HCCs with a count of payment HCCs, with little loss of predictive accuracy. As a result, we decided to use a single interaction variable to indicate "severe" illness, defined as the presence of at least one severe illness HCC for an enrollee, with binary variables for the adult enrollee's total number of payment HCCs = 1, 2, 3, ..., 10+. The proposed interacted HCC counts approach applies slightly differently in the child models, where the severe illness HCC count factors apply from 1, 2, 3, ..., to 8+ payment HCCs.

We also considered including 8 organ transplant status HCCs (heart, lung, liver, etc. transplants) among our severe illness HCCs. We added a set of variables for transplant status interacted with total HCC count to boost predicted expenditures even further for organ transplant enrollees with a large number of HCCs. After some experimentation, we determined that interacting transplant status with number of payment HCCs = 4, 5, 6, 7, 8+ in the adult models produced good predictive accuracy and had sufficient sample sizes for reasonably stable factor estimates. We found that organ transplant adult enrollees with fewer than 4 payment HCCs were very rare, and that including those interactions did not improve predictive accuracy for the highest-risk enrollees. For the child models, having one variable for 4+ payment HCCs provided more stable estimates given the smaller sample sizes for children than for adults. Table 4.4 provides an overview on the proposed interacted HCC count factors for both the child and adult models.

Table 4.4 Structure of the Interacted HCC Count Factors by Risk Adjustment Model Type

	Severity HCC Counts	Transplant HCC Counts
Adult Model Factors	1, 2, 3, 4, 5, 6, 7, 8, 9, 10+	4, 5, 6, 7, 8+
Child Model Factors	1, 2, 3, 4, 5, 6, 7, 8+	4+

When we were investigating the various HCC count approaches, incorporating safeguards to protect against the potential for gaming was a major consideration. We note that given individual HCCs are already in the risk adjustment model, the incentive to code for more HCCs already exists. However, we were specifically concerned that the presence of counts across all HCCs may further incentivize issuers to code for more HCCs, thus increasing their payment or reducing their charge under the state payment transfer formula. This would be inconsistent with the risk adjustment principle not to encourage coding proliferation.[93] However,

[92] HCC 18 (Pancreas Transplant) was not included in the list of HCCs selected for the Proposed Interacted HCC Counts Variables for the Adult and Child Models in the proposed 2022 Payment Notice. *See* Table 3 in the proposed 2022 Payment Notice, 85 FR at 78593. This was due to the fact that HCC 18 has a much lower coefficient than the other transplant HCCs in the adult models and was not underpredicted by the models. We include it in the analysis in this paper to be consistent with the other transplant HCCs for purposes of the further evaluation of this model specification change.

[93] See Section 1.1.2 for information on the principles that guide the HHS risk adjustment models' diagnostic classification system and, in particular, Principle 6: The diagnostic classification should not reward coding proliferation.

we believe that using the proposed interacted HCC count approach, which restricts the incremental risk score adjustment to enrollees with at least one severe illness HCC, lessens the concerns that issuers may attempt to inflate HCC counts to influence their transfers under the state payment transfer formula. Our analysis of the 2016, 2017, and 2018 enrollee-level EDGE datasets revealed that severe illness HCCs are relatively uncommon; less than 2 percent of the adult enrollee-level EDGE data population across these three benefit years had at least one severe illness HCC, as opposed to about 20 percent of adult enrollees with any payment HCC. Therefore, the scope for potentially inflating HCC coding frequency would be limited to a small fraction of total enrollees.

4.3 How the Interacted HCC Counts Approach Works

Based on our above analyses, we proposed in the proposed 2022 Payment Notice[94] to drop the single severity indicator factors in the HHS risk adjustment adult models and replace them with two sets of "interacted HCC counts" variables in the adult and child models, a set of 10 factors triggered by the presence of at least one of an expanded list of severe illness HCCs, and 5 parameters triggered by the presence of at least one transplant HCC, and differentiated by the enrollee's total count of HCCs. Appendix C lists the severe illness and transplant HCCs that were proposed as part of the interacted HCC counts approach in the proposed 2022 Payment Notice.[95] While adding the interacted HCC counts approach to the risk adjustment adult and child models introduces some additional complexity, this is offset by a reduction in complexity due to the removal of the existing severity indicator factors in the adult models. In addition, this change would improve the prediction of the very highest-risk enrollees.

Under the interacted HCC counts approach, when an enrollee has a severe illness HCC, the enrollee's risk score adjustment includes the applicable severe illness HCC counts variable factor, in addition to the applicable HCC factor. The proposed interacted HCC counts factors are based on the counts of all payment HCCs for an enrollee with at least one HCC that is classified as a severe illness or transplant HCC. For example, if an enrollee has diabetes, sepsis, heart failure, and asthma (with sepsis being a severe illness HCC), the enrollee would get credit for all the HCCs listed, plus an additional amount determined based on the total count of all payment HCCs for the enrollee, which in this case would be 4.

Tables 4.5.1 and 4.5.2 show the interacted HCC counts factors for severe illness and transplant status that were proposed in the proposed 2022 Payment Notice based on the number of payment HCC counts for enrollees in the adult silver model. When looking at the interacted HCC counts factors, it is important to consider them in tandem with the enrollee's underlying HCC factors. Even though most of the proposed severe illness HCC count factors are negative, adding a severe illness HCC always increases the enrollee's risk score, because the sum of the severe illness HCC and the interacted HCC count factors is always positive. For example, under this approach, the proposed adult silver model factor in the proposed 2022 Payment Notice for Viral or Unspecified Meningitis (HCC 4), which was proposed as a severe illness HCC, is 8.323; when combined with the smallest severe illness factor for one HCC of -6.181, this diagnosis

[94] 85 FR 78583.
[95] See supra note 92 on HCC selection.

would increase the enrollee's risk score by 2.142. Moreover, an increase in the count of HCCs leads to a monotonic increase in the enrollee risk score, because the severe illness factors are less negative (and sometimes positive) with a larger number of HCCs.

Table 4.5.1 Adult Silver Plan Model Severe Illness Count Factors Proposed in the Proposed 2022 Payment Notice

Severe Illness Factor	
Severe illness, 1 payment HCC	-6.181
Severe illness, 2 payment HCCs	-5.824
Severe illness, 3 payment HCCs	-4.526
Severe illness, 4 payment HCCs	-3.415
Severe illness, 5 payment HCCs	-2.554
Severe illness, 6 payment HCCs	-2.103
Severe illness, 7 payment HCCs	-0.987
Severe illness, 8 payment HCCs	-0.328
Severe illness, 9 payment HCCs	1.458
Severe illness, 10+ payment HCCs	10.431

Table 4.5.2 Adult Silver Plan Model Transplant Count Factors Proposed in the Proposed 2022 Payment Notice

Transplant Status Factor	
Transplant, 4 payment HCCs	3.483
Transplant, 5 payment HCCs	7.353
Transplant, 6 payment HCCs	12.622
Transplant, 7 payment HCCs	18.688
Transplant, 8+ payment HCCs	33.829

To illustrate an example of the potential interacted HCC counts approach, Table 4.6.1 shows a hypothetical 63-year-old male, enrolled for a full 12 months in a silver plan, with diabetes without complication (HCC 21), which is not a severe illness HCC, and no other risk markers. Under the interacted HCC counts approach, his predicted cost is 0.605 (based on the adult model silver plan factors published in the proposed 2022 Payment Notice). Predicted costs are relative to expected costs for the average enrollee; therefore, an enrollee with a predicted cost of 0.605 is expected to be roughly 60 percent as costly as the average enrollee.

Table 4.6.1 Sample Risk Score Calculation: 1 Payment HCC (No Severe Illness HCC)

Age-Sex	HCC	Severe Illness		Predicted Cost
63-year-old male	21	No		
0.343	0.262		=	0.605

If this enrollee develops sepsis, he would also get the factor for HCC 2 (Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock), which is a severe illness HCC, and the associated interacted HCC counts factor for two HCCs, which results in his predicted cost increasing to 4.175, as shown in Table 4.6.2. As previously stated, it is important to view the severe illness HCC count factor together with the enrollee's underlying HCCs. Even though the severe illness HCC count factor for two HCCs is negative (-5.824), when combined with the factor for the severe illness HCC sepsis (9.394), the result is an increase in the enrollee's predicted cost.

Table 4.6.2 Sample Risk Score Calculation: 2 Payment HCCs (With Severe Illness HCC)

Age-Sex	HCC	HCC	Severe Illness	Severe Illness Count Factor		Predicted Cost
63-year-old male	21	2	Yes	2		
0.343	0.262	9.394		-5.824	=	4.175

If this enrollee also develops heart failure, which is not a severe illness HCC, along with the two previously discussed HCCs, he would also get the factor for HCC 130 (Heart Failure) and the enrollee's predicted cost would rise to 7.347, as shown in Table 4.6.3.

Table 4.6.3 Sample Risk Score Calculation: 3 Payment HCCs (With Severe Illness HCC)

Age-Sex	HCC	HCC	HCC	Severe Illness	Severe Illness Count Factor		Predicted Cost
63-year-old male	21	2	130	Yes	3		
0.343	0.262	9.394	1.874		-4.526	=	7.347

The purpose of adding severity and transplant indicators interacted with HCC counts factors is to account for the fact that costs of certain HCCs rise significantly when they occur with multiple other HCCs. Although often negative, the values of the proposed severe illness factors rise with the enrollee's total number of HCCs, increasing the enrollee's total predicted cost as the number of HCCs increases. As seen in the tables above, when the enrollee with a severe illness HCC (sepsis) in addition to diabetes acquires the heart failure HCC, the total count of payment HCCs increases from 2 to 3. As a result, this illustrative enrollee's predicted cost without heart failure (4.175) rises by 3.172, which is equal to the incremental cost estimated for heart failure (1.874) *plus* the increase in the severe illness HCC count from 2 HCCs to 3 HCCs (-4.526 – (-5.824) = 1.298). This approach aligns with risk adjustment principles that the diagnostic classification should encourage specific coding and that providers and issuers should not be penalized for coding additional conditions.[96]

The proposed interacted HCC counts approach also makes an adjustment for the costs of enrollees with a transplant status who have 4 or more payment HCCs. If our above hypothetical 63-year-old male enrollee with sepsis, diabetes, and heart failure also had kidney transplant status (HCC 183), in addition to his other conditions and the interacted HCC count factors for severe illness, he would also get an interacted HCC counts factor for his transplant, resulting in his predicted cost increasing from 7.347 to 20.444 as shown in Table 4.6.4.

Table 4.6.4 Sample Risk Score Calculation: 4 Payment HCCs (With Severe Illness HCC, Transplant HCC)

Age-Sex	HCC	HCC	HCC	HCC	Severe Illness	Transplant	Severe Illness Count Factor	Transplant Count Factor		Predicted Cost
63-year-old	21	2	130	183	Yes	Yes	4	4		
0.343	0.262	9.394	1.874	8.503			-3.415	3.4383	=	20.444

For an enrollee with a transplant HCC, the incremental transplant HCC add-on factor applies if the total number of payment HCCs is greater than or equal to 4. Under the proposed interacted HCC counts approach, transplant HCCs are a subset of severe illness HCCs. Therefore, an enrollee with a transplant HCC and at least 4 HCCs will have both the severe illness HCC add-on and the transplant add-on included as part of the enrollee's risk score calculation. If the enrollee's total number of payment HCCs was less than 4, they would not receive any incremental transplant add-on, but they would receive the incremental severe illness add-on.

The proposed severe illness adjustment in the interacted HCC counts approach proposed in the 2022 Payment Notice increases with an enrollee's total disease burden as measured by number of HCCs for enrollees with a severe illness or transplant HCC. Therefore, the proposed interacted HCC counts approach in the proposed 2022 Payment Notice allows predicted costs to

[96] See Section 1.1.2 for information on the principles that guide the HHS risk adjustment models' diagnostic classification system and, in particular, Principle 5: The diagnostic classification should encourage specific coding and Principle 7: Providers should not be penalized for recording additional diagnoses (monotonicity).

increase for the highest-risk, most expensive enrollees, who tend to have one or more severe illness HCCs plus a large number of other payment HCCs. The current adult models' severe illness adjustment, by contrast, is a single fixed increment that applies to fewer enrollees, both because fewer HCCs are classified as "severe illness" under the models, and because the adjustment is only made when other selected HCCs are present in addition to the severe illness HCC. Based on our review and analysis of the 2018 enrollee-level EDGE dataset, the current models' adjustment applied to 45,931 enrollees, or 0.24 percent of adult enrollees, while the proposed 2022 Payment Notice approach would apply to 258,755 enrollees, or 1.34 percent of adult enrollees.

Figure 4.1 shows risk score changes for selected severe illness HCCs based on the number of HCCs accounting for all the proposed model specification changes in the proposed 2022 Payment Notice (i.e., two-stage weighting, interacted HCC counts, and HCC-contingent enrollment duration factors). As Figure 4.1 demonstrates, there is a consistent quantitative pattern for all of the selected severe illness HCCs where the risk score contribution increases as the counts of payment HCCs increase. This effect is largely due to the interacted HCC counts approach where an enrollee's severe illness adjustment increases as an enrollee's total number of payment HCCs increases.

Figures 4.1: Proposed 2022 Payment Notice Model Specification Updates Impacts on Selected HCCs[97]

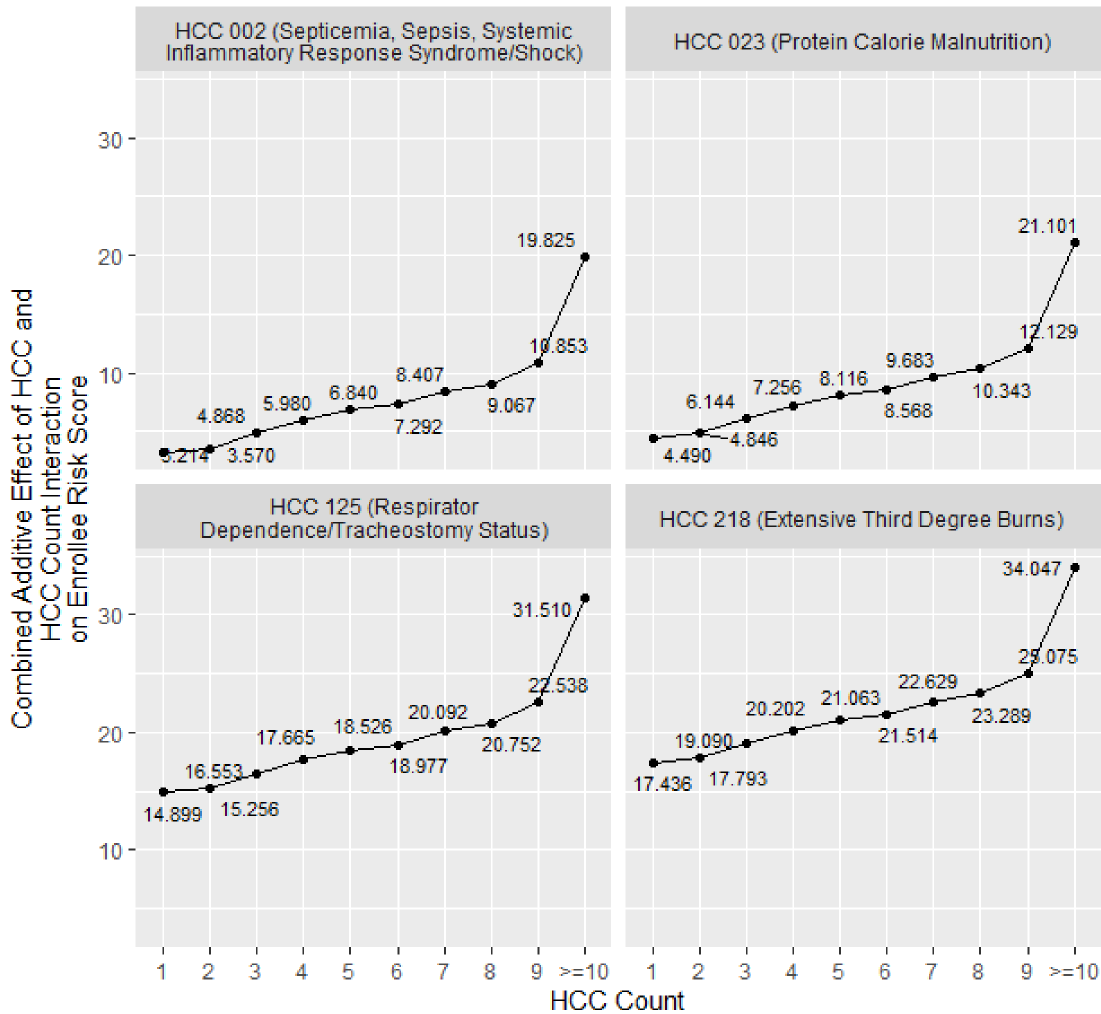


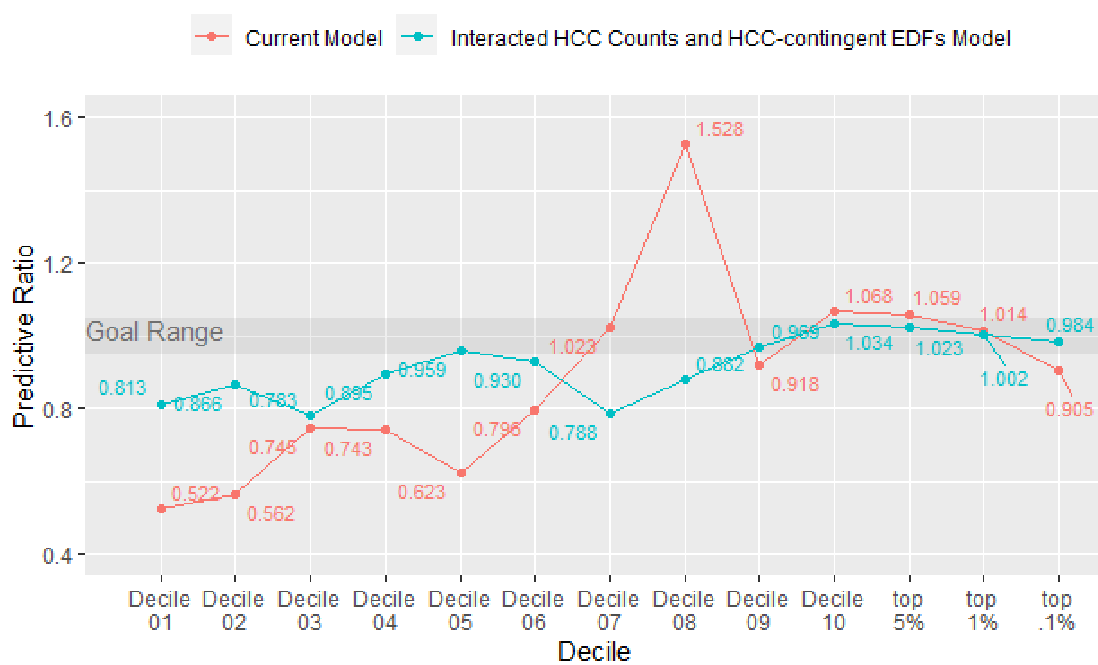
Figure 4.1 does not represent full risk score calculations, instead it reflects only the factors relevant for the selected severe illness HCCs. For example, the risk score contribution for sepsis at 4 HCCs shown in the first panel of Figure 4.1 includes the factor for sepsis plus the interacted HCC count add-on factor for 4 payment HCCs, but does not include the individual factors for the 3 other HCCs that the enrollee would have, or other factors such as age, sex, or RXCs. As Figure 4.1 shows, the inclusion of each additional HCC to the count increases the interacted HCC count add-on factor and results in an increase to the add-on factor's contribution to the risk score.

[97] The data used in Figure 4.1 uses blended factors (i.e. averages across multiple data years, 2016, 2017, and 2018). The rest of Chapter 4 focuses on model factors (and, more often, PRs) associated with a single year of factors.

4.4 Interacted HCC Counts Approach Change Impact

The purpose of this section is to describe the impact of the potential introduction of interacted HCC counts factors into the current adult and child models. One of the benefits of the proposed interacted HCC counts approach is that it improves prediction for both the lowest- and highest-risk enrollees compared to the current models. Figure 4.2 shows the PRs by decile of predicted plan liability in the adult silver plan model using the 2018 enrollee-level EDGE dataset. The proposed interacted HCC counts approach (combined with the HCC-contingent enrollment duration factors, discussed in Chapter 3)[98] allows for the risk adjustment models to better account for the higher costs of enrollees with certain high-cost HCCs (i.e. severe illness and transplants). At the same time, differentiating better between low- and high-cost enrollees enables the models to predict costs more accurately for lower-risk enrollees. This results in improved prediction for both the lowest-risk and very highest-risk enrollees compared with the current adult models, even without accounting for the proposed two-staged weighted approach discussed in Chapter 2. For example, our review of the 2018 enrollee-level EDGE dataset found the proposed interacted HCC counts approach combined with the proposed HCC-contingent enrollment duration factors improves the PR for adult enrollees in risk decile 1 from 0.52 to 0.81 and the PR for adult enrollees without HCCs from 0.79 to 0.90.

Figure 4.2: Adult Silver Plan Model Predictive Ratios by Decile, 2018 Enrollee-Level EDGE Dataset

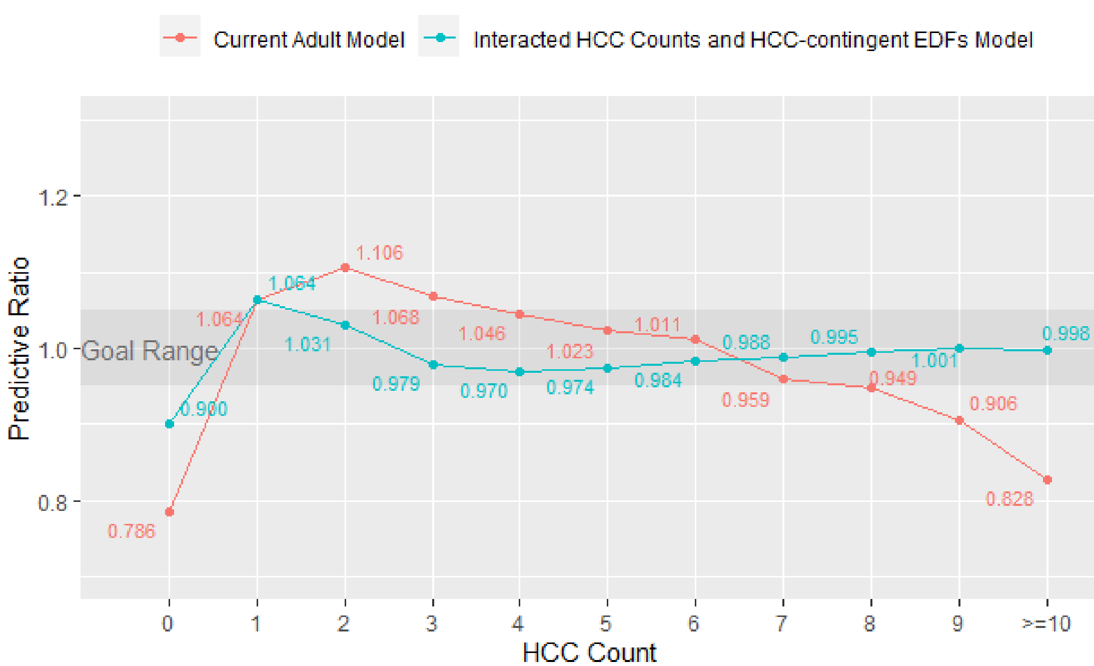


[98] During testing of the different model approaches, we combined the proposed 2022 Payment Notice enrollment duration factor change into our model testing of the interacted HCC counts factors to reduce the number of model types that we were testing at once and to ensure that the results of the interacted HCC count factors would already take into account the reduced overprediction of enrollee costs at decile 8.

In particular, we found that the adoption of the proposed interacted HCC counts and proposed HCC-contingent enrollment duration factors approach in the adult models would improve prediction for enrollees at the highest percentiles of plan liability, particularly in the 10th decile, 5%, 1%, and especially 0.1%. For example, using the adult silver plan model in the 2018 enrollee-level EDGE dataset, the adoption of the proposed interacted HCC counts and proposed HCC-contingent enrollment duration factors improves the PR for the top 0.1 percent of adult silver plan enrollees from 0.91 to 0.98 and the PR for enrollees with 10 or more HCCs from 0.83 to 1.00. We saw similar improvements in the adult models using the 2016 and 2017 enrollee-level EDGE datasets. The adoption of the proposed interacted HCC counts approach[99] also shows improvements in the child models using the 2016, 2017, and 2018 enrollee-level EDGE datasets.

Figure 4.3 shows the PRs by number of HCCs in the adult silver plan model using the 2018 enrollee-level EDGE dataset. As with the PRs by decile in Figure 4.2, the proposed interacted HCC counts approach combined with the proposed HCC-contingent enrollment duration factors significantly improves prediction for adult enrollees across most HCC counts. This is particularly the case for adult enrollees with high numbers of HCCs (greater than 6). We saw similar results for the adult models using the 2016 and 2017 enrollee-level EDGE datasets, as well as when analyzing the impact of the proposed interacted HCC counts approach on the child models.[100]

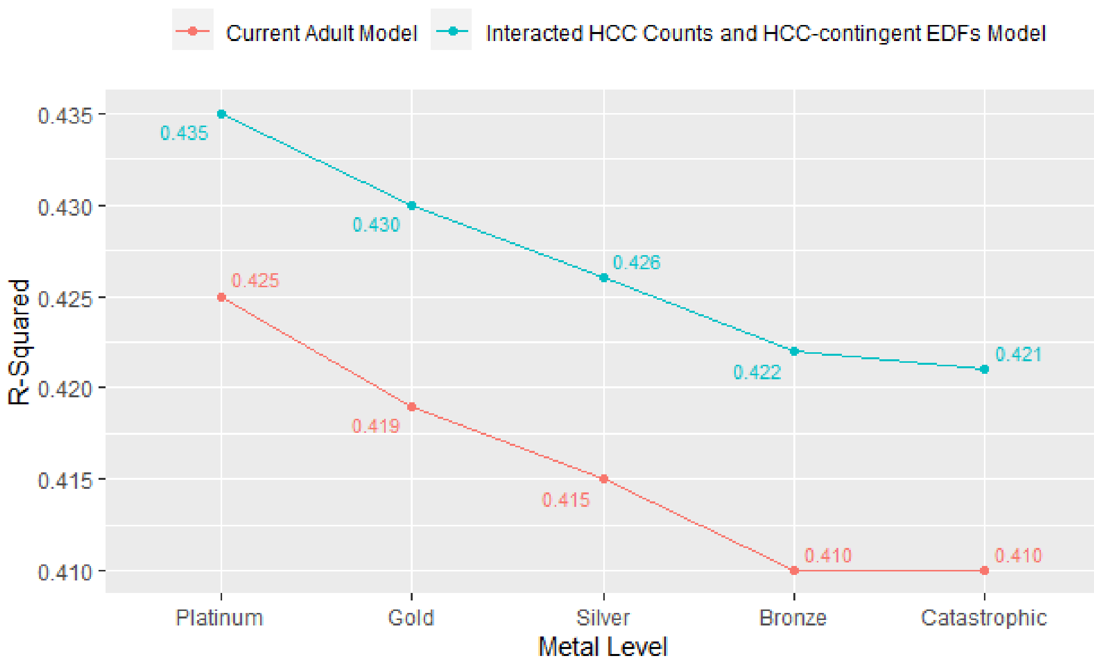
Figure 4.3: Adult Silver Plan Model Predictive Ratios by Number of HCCs, 2018 Enrollee-Level EDGE Dataset



[99] As discussed above, the current child models do not include enrollment duration factors and the proposed HCC contingent enrollment duration factors would only apply to the adult models. Our analysis of the child models therefore focused on the impact of the proposed interacted HCC counts approach.
[100] Ibid.

The proposed interacted HCC counts and proposed HCC-contingent enrollment duration factors approach also demonstrated improved R-squared statistics. Figure 4.4 shows a comparison for the 2018 benefit year of the R-squared statistics for the current adult models and the proposed interacted HCC counts approach combined with the HCC-contingent enrollment duration factors for the adult models. As seen in Figure 4.4, the combination of these proposals improves the R-squared statistics across all metal levels in the adult models. Our analysis showed similar improvements in the child models when the proposed interacted HCC counts approach is applied.[101]

Figure 4.4: R-Squared Statistics for the Adult Models, 2018 Enrollee-Level EDGE Dataset



One benefit of the proposed interacted HCC count approach is that it would not overhaul the existing risk adjustment factors and would build upon the current models. As discussed in Chapter 2, while other approaches we previously considered (such as the non-linear and full HCC counts approaches) would result in factors that would be substantially different than under the current models, under the proposed interacted HCC count approach, the factors would remain fairly stable compared to the current factors and models. In addition, the proposed interacted HCC counts approach can be used in combination with other refinements, such as the proposed two-stage weighted approach and the proposed HCC-contingent enrollment duration factors, and can be easily modified, adjusted, expanded, or constrained in the future to include additional HCCs or to remove HCCs. As previously mentioned, another advantage the proposed interacted HCC counts approach has over the other approaches we considered is that it uses relatively uncommon HCCs, which we believe will mitigate concerns about potentially gaming through reporting of more HCCs. More specifically, by limiting the proposed interacted HCC counts factors to certain severe illness and transplant HCCs, we believe that the proposed interacted

[101] See supra note 990.

HCC counts factors would restrict the scope for coding proliferation and effectively mitigate the potential for gaming.

One potential concern is that the interacted HCC counts coefficients might be based on small sample sizes. To address these concerns, we considered sample sizes of the various interacted HCC counts factors when developing the factors proposed in the proposed 2022 Payment Notice. To that end, we analyzed multiple years of enrollee-level EDGE datasets and chose the model specifications that grouped the HCC counts interacted with individual severity and transplant HCCs into two sets of aggregated factors to maximize sample size, reduce concerns of overfitting the model, and reduce the number of factors being added to the models. The resulting sample sizes for the proposed interacted HCC counts approach in the proposed 2022 Payment Notice were consistent with the sample sizes for individual HCCs in the current adult and child risk adjustment models.

4.5 Conclusion

In summary, the proposed interacted HCC counts approach significantly improves predictions across most deciles and HCC counts for the very highest-risk enrollees, as well as lowest-risk enrollees without HCCs. As discussed previously, we tested the impact of the interacted HCC counts model approach with the HCC-contingent enrollment duration factors to limit the number of model options that we were comparing in our analysis and to ensure that the impact of the reduced overprediction of enrollee costs at decile 8 (due to the HCC-contingent enrollment duration factors) was accounted for when considering the impact of the proposed interacted HCC counts approach. Our analysis of the combined refinements did not worsen model performance for any of the subpopulations examined, and the overall model R-squared rises. However, while the combined refinements showed improvement for lowest-risk enrollees, such improvement is greatly increased when the combined with the two-stage weighted approach discussed in Chapter 2. The proposed interacted HCC counts, HCC-contingent enrollment duration factors, and the two-stage weighting, when combined, can optimize model performance across the risk spectrum (i.e., for lowest-risk, medium risk and very highest-risk enrollees), as demonstrated in Chapter 5.

Chapter 5: Results and Estimated Impact of Model Specification Changes Under Consideration

In response to comments received during the 2022 Payment Notice rulemaking process on the proposed model specification changes described in Chapters 2-4 of this paper, HHS plans to separately release transfer estimates from a simulation conducted using data on issuers' EDGE servers, described in section 5.2, towards the end of 2021. The purpose of providing the information in this chapter and simulated transfer estimates at a later date is to give issuers a further opportunity to assess how these proposed risk adjustment model specification changes may affect risk scores and transfers under the state payment transfer formula.[102]

For this chapter, we first describe model performance metrics of the combined proposed model specification changes using the 2018 enrollee-level EDGE dataset. Then, we outline the methodology behind the transfer simulation that is being conducted using data on issuers' EDGE servers to provide simulated issuer-specific results. This simulation is applying the proposed 2022 model specification changes to 2020 benefit year data from issuers' EDGE servers to estimate what 2020 benefit year transfers would have been if the proposed 2022 model specification changes were applied. We selected 2020 benefit year data for this exercise because it was the most recently available data at the time of conducting the analysis.

5.1 Combined Impact of Model Specification Changes

We presented the estimated impact of the two-stage weighted approach on model performance in Chapter 2, and the estimated impact of the HCC-contingent enrollment duration factors and the interacted HCC counts on model performance in Chapters 3 and 4, respectively. These impacts illustrate two ways in which the model specification changes proposed in the proposed 2022 Payment Notice could be partially implemented. In the proposed 2022 Payment Notice, we considered the potential trade-offs of using the proposed model specifications changes (i.e., two-stage weighted approach, HCC-contingent enrollment duration factors, and interacted HCC counts approach) over the current model specifications. We also explored how to balance the goals of improving the models' prediction with mitigating modeling complexity and gaming concerns, along with promoting market stability.

HHS continues to believe the best way to comprehensively improve the predictive accuracy of the models across the risk spectrum is to implement all three model specification changes together. This section focuses on the combined impact of all the model specification changes described in Chapters 2–4 on model performance, as compared to the current models.

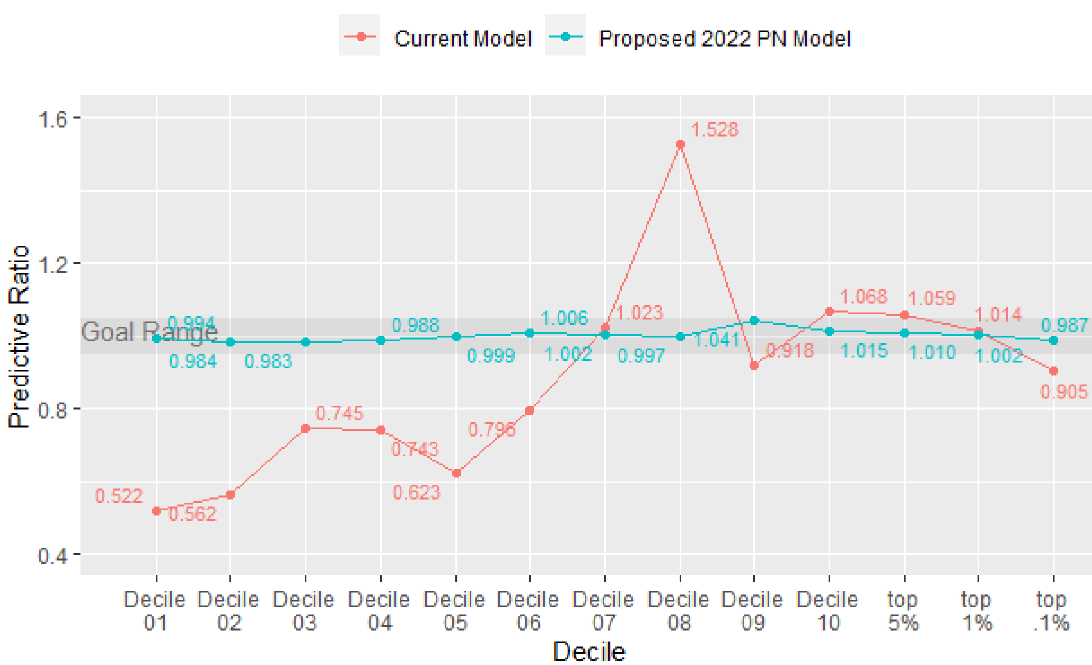
The following figures compare model performance between the current adult models and the combined proposed model specification changes as described in Chapters 2–4 of this paper, as applied to the 2018 enrollee-level EDGE dataset for adult enrollees that include:

[102] If an issuer wishes to use the simulation results to assist in assessing the impact of the model specification changes on future benefit year transfer amounts, it should do so with caution and in combination with other significant data. In particular, smaller issuers may experience a wider degree of variation, given the impact that larger issuers have on transfers within a state market risk pool.

1) The two-stage weighted approach that reweights lowest-risk enrollees more heavily,
2) Monthly enrollment duration factors of up to 6 months for only partial-year enrollees with HCCs, and
3) The interacted HCC counts approach to account for additional incremental risk for enrollees with multiple conditions.

We found that, together, these changes improved PRs in comparison to the current models in each decile of predicted plan liability. Figure 5.1 shows PRs by decile of predicted plan liability for the adult silver plan model using the 2018 enrollee-level EDGE dataset. Additionally, unlike the current adult models (see Figure 2.1), the PRs by age-sex factor for silver plan adult enrollees with and without HCCs are mostly within an acceptable range close to 1.0 when these combined proposed model specification changes are applied, as seen in Figure 5.2. We found similar patterns for PRs by decile of predicted plan liability and age-sex factor across all metal levels for the adult models.

Figure 5.1 Adult Silver Plan Model Predictive Ratios by Decile of Predicted Plan Liability, 2018 Enrollee-Level EDGE Dataset[103]



[103] As previously mentioned, "Enrollee-Level EDGE Dataset" in this paper refers to the national dataset used for risk adjustment model recalibration. This dataset contains limited information generated by reports HHS receives from issuers' EDGE servers, and excludes plan, state, and other information necessary for calculating risk adjustment transfers.

Figure 5.2 Distribution of Proposed 2022 Payment Notice Adult Silver Plan Model Predictive Ratios by Age-Sex Factor, 2018 Enrollee-Level EDGE Dataset

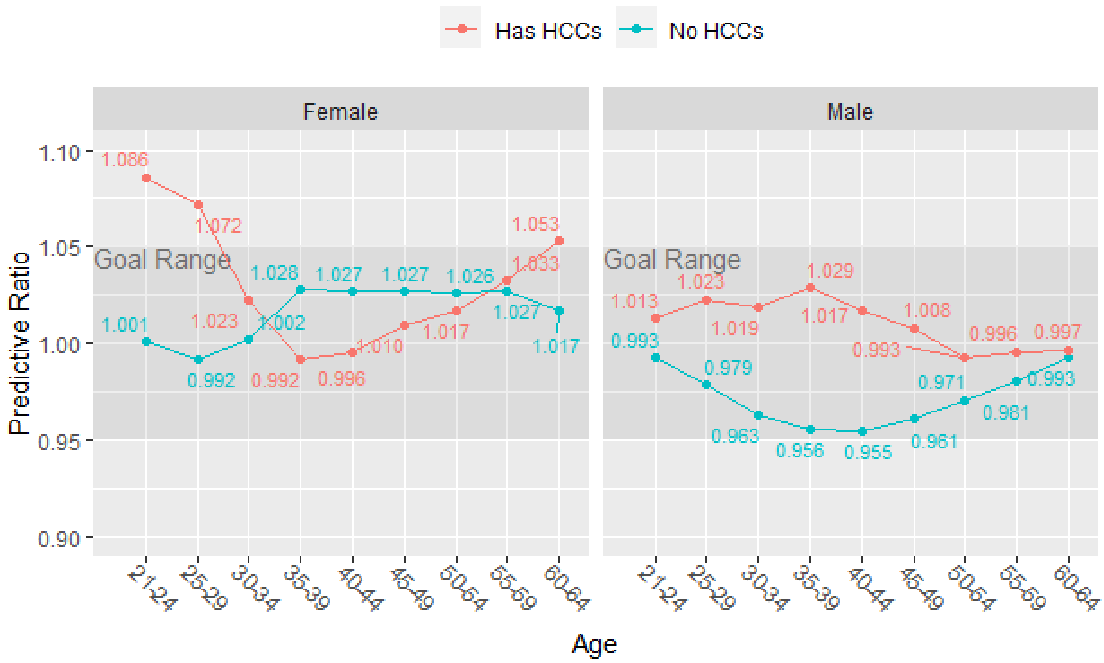


Figure 5.3 shows the adult silver plan model PRs by enrollment length for the current adult models and the proposed 2022 Payment Notice models. When compared to the current model with HCC-contingent EDFs only, adding the two-stage weighting approach and interacted HCC counts model specification weakens the impact of the HCC-contingent enrollment duration factors. The PRs for the models with only HCC-contingent EDFs are slightly closer to 1.0 because the variables for number of HCCs and enrollment duration are explicitly included in the model with HCC-contingent EDFs only. Re-weighting enrollees with lower predicted plan liability during two-stage weighting can cause the PRs for factors in the model to deviate more from 1.0. While some underprediction still remains for enrollees with HCCs and shorter enrollment lengths when the proposed model specification changes are combined, as seen in Figure 5.3, the underprediction is substantially reduced compared to the current models. Similarly, for enrollees without HCCs and shorter enrollment lengths, predictive accuracy improves under the combined proposed model specification changes compared to the current models.

Figure 5.3 Adult Silver Plan Model Predictive Ratios by Enrollment Length for Current Model, Current Model with HCC-Contingent EDFs, and Proposed 2022 Payment Notice Model, 2018 Enrollee-Level EDGE Dataset

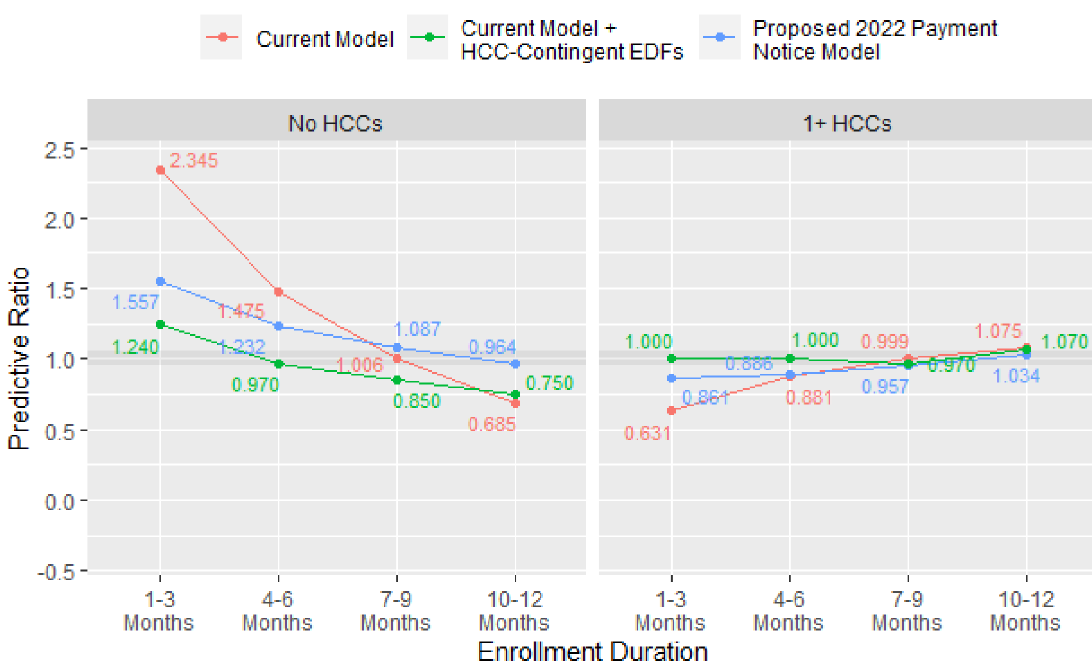
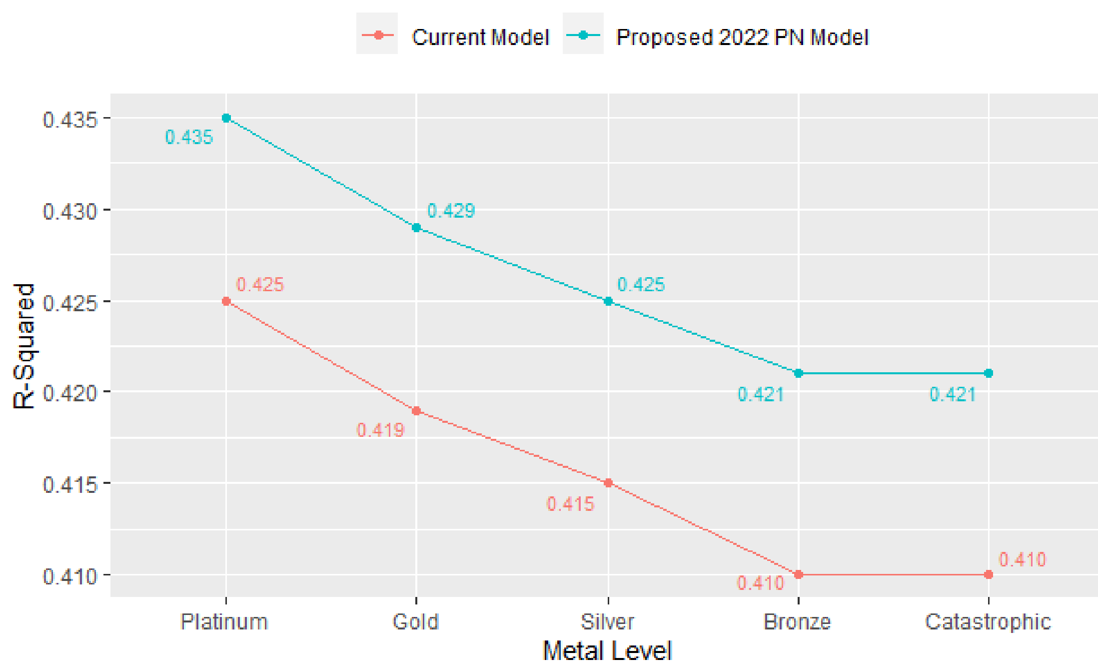


Figure 5.4 shows the R-squared statistics for the current adult models compared to the proposed 2022 Payment Notice models, using the 2018 enrollee-level EDGE dataset. The proposed 2022 Payment Notice adult models have higher R-squared statistics than the current models, indicating a better individual-level fit.

Figure 5.4 Adult Model R-Squared Statistics, 2018 Enrollee-Level EDGE Dataset[104]



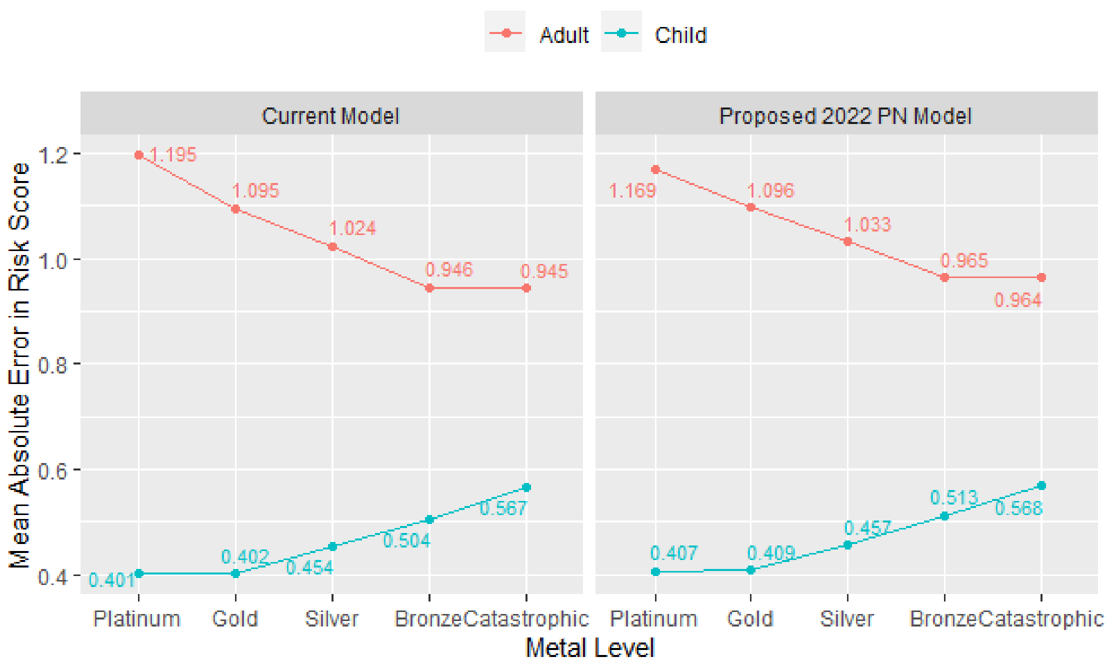
Though we considered PRs and, to a lesser extent, R-squared statistics to be our primary performance metrics, we also considered mean absolute error (MAE) as an alternative metric to evaluate the impact of the model specification changes described in Chapters 2–4. MAEs measure the closeness of the prediction to the eventual outcomes and, unlike R-squared, which measures the proportion of the variance in predicted values explained by the model factors, depend on the absolute magnitude of the data itself. For example, compared to the adult model MAEs, the child model MAEs are lower overall because mean costs are lower and there is less cost variation among children. A low MAE indicates small prediction errors, so all else equal, lower values are more desirable.

We found that the MAEs did not materially differ between the current adult and child models and the proposed 2022 Payment Notice models, which incorporated the combined proposed model specification changes. Figure 5.5 shows the MAEs for the current adult and child models versus the proposed 2022 Payment Notice models. Compared to the current adult models, the proposed 2022 Payment Notice adult models have very slightly higher MAEs for all metal levels except for platinum. This is driven by the two-stage weighted approach, which introduces more variability by weighting observations with lower predicted plan liability higher

[104] For each of the HHS risk adjustment current and proposed models, the R-squared statistic is in the range of published estimates for concurrent risk adjustment models. See Hileman, Geof and Spenser Steele. "Accuracy of Claims-Based Risk Scoring Models." Society of Actuaries. October 2016.

instead of predicting overall plan liability for all enrollees. We found negligible differences between the current and proposed 2022 Payment Notice child models. Overall, the improvements in predictive ratios and R-squared between the current child models and the proposed child models (reflecting the two-stage weighted approach and the interacted HCC counts approach) are not associated with any meaningful increase in MAE.

Figure 5.5 Adult and Child Models' Mean Absolute Errors, 2018 Enrollee-Level EDGE Dataset



After evaluating model performance metrics of the combined proposed model specification changes, we found that, together, these changes are expected to improve model performance in comparison to the current models. Our analysis found general improvement in PRs for the models with the combined proposed model specification changes across each decile of predicted plan liability, by age-sex factor for adult enrollees with and without HCCs, and by enrollment length. We also observed higher R-squared squared statistics across metal levels and similar MAEs compared to the current models. These observations support our belief that the best way to comprehensively improve the predictive accuracy of the models across the risk spectrum is to implement all three model specification changes described in Chapters 2-4 together.

5.2 EDGE Server Transfer Simulation

To give issuers additional information they can use to assess the impact of the combined proposed model specification changes described in Chapters 2–4 on transfer results under the state payment transfer formula, HHS is conducting a transfer simulation on 2020 benefit year plan-level data extracted from issuers' EDGE servers and providing issuer-specific risk score and transfer estimates.[105] This section provides an overview of the methodology and operational process for running this transfer simulation and explains what issuers can expect from this transfer simulation. We intend to release the transfer simulation results using 2020 plan-level EDGE data towards the end of 2021.

Given the availability of 2020 benefit year data on issuers' EDGE servers, we are using it to conduct this transfer simulation and provide comparative transfer simulation results later this year. Since the purpose of this transfer simulation is to estimate the change in transfers due to model specification changes, the underlying benefit year of data used for the simulation should not matter.

To estimate the impact of the proposed model specification changes, HHS is simulating the process used to calculate transfers for the applicable (un-simulated) benefit year. HHS will first calculate what transfers would be under the current models that do not reflect the combined proposed model specification changes described in Chapters 2 – 4. To provide this benchmark, HHS will use the 2020 plan-level EDGE data to simulate what transfers would be under the most current risk adjustment model factors, the final 2022 risk adjustment model factors.[106] HHS will then use the 2020 plan-level EDGE data to simulate transfers under the risk adjustment model factors as they were proposed in the proposed 2022 Payment Notice, which reflected the adoption of the combined proposed model specifications changes described in Chapters 2 – 4.[107] Stakeholders will be able to compare the results of these two simulations to further analyze the potential impact of the proposed model specification changes.

For the simulation to run successfully, issuers need to run multiple EDGE Ad Hoc commands on their respective EDGE servers.[108] To support this, HHS therefore needs to:

Step 1: Send EDGE server commands to have issuers copy final 2020 benefit year risk adjustment data, so that it can be used separately for this transfer simulation exercise;

Step 2: Send EDGE server commands to apply the final 2022 model factors to the copied 2020 data on issuers' EDGE servers;

---

[105] If an issuer wishes to use the simulation results to assist in assessing the impact of the model specification changes on future benefit year transfer amounts, it should do so with caution and in combination with other significant data.

[106] See Updated 2022 Benefit Year Final HHS Risk Adjustment Model Coefficients at: https://www.cms.gov/files/document/updated-2022-benefit-year-final-hhs-risk-adjustment-model-coefficients-clean-version-508.pdf.

[107] See 85 FR at 78586 – 78596.

[108] Issuers are critical partners for running this simulation and need to participate by running the commands provided by CMS within the designated timeframes in order to receive the issuer-level simulation reports. If a credible issuer within a state does not participate, HHS would also not be able to provide simulation results for that state market risk pool. Issuers are considered credible if they have at least 0.5 percent market share in a state. Information on the commands and accompanying timeframes can be found on REGTAP at: https://www.regtap.info/reg_library.php?libfilter_topic=3&libfilter_keyword=xtestyz.

Step 3: Use plan-level EDGE data extracted from issuers' EDGE servers from Step 2 to estimate transfer amounts under the final 2022 models that do not reflect the combined proposed model specification changes;

Step 4: Send EDGE server commands to apply the model factors as they were proposed in the proposed 2022 Payment Notice to the copied 2020 data on issuers' EDGE servers;

Step 5: Use plan-level EDGE data extracted from issuers' EDGE servers from Step 4 to estimate transfer amounts under the proposed 2022 Payment Notice models that reflect the proposed model specification changes; and

Step 6: Compare the transfer simulation results from Steps 3 and 5 to provide further analysis of the impact of the proposed model specification changes.

Table 5.5 below shows the timeline of EDGE server simulation (SIM) zone command iterations. There are four iterations of the SIM zone command where issuers must execute commands as soon as HHS deploys them. Iterations begin on the following dates only for issuers that have successfully completed earlier iterations. For each iteration, separate commands are being released on various dates to:

1. Create simulation zone schema
2. Run the Enrollee Claims Summary (ECS), Frequency (FREQ), Risk Adjustment (RA), High Cost Risk Pool (HCRP), and Ad Hoc commands
3. Archive simulation zone schema

Following the schedule below, issuers have already begun running SIM zone commands on their EDGE servers to support this simulation. We intend to release the transfer simulation results towards the end of 2021, once issuers finish running SIM zone commands, plan-level EDGE data is extracted from issuers' EDGE servers to calculate transfers under the state payment transfer formula, and CMS analyzes the results.

Table 5.5 EDGE Server Simulation (SIM) Zone Commands

Iteration	Tentative Start Date	Description
1	07/16/21	• Data: Final BY20 Data and BY20 Reference Data (RDV42) • CMS deploys commands including Create SIM, ECS, FREQ, RA, HCRP, Ad Hoc, and Archive SIM commands • Issuer and HHS receive reports for remote commands • Note: Issuers who have already copied their BY20 production (PROD) data to the SIM zone for discrepancy analysis and/or beta testing will not receive the 'Iteration 1' Copy PROD to SIM command.
2	07/31/21	• Data: Final BY20 Data and BY21 DIY Table 9 Data • CMS deploys commands including Create SIM, ECS, FREQ, RA, HCRP, Ad Hoc, and Archive SIM commands • Issuer and HHS receive reports

3	08/21/21	• Data: Final BY20 Data and Final BY22 DIY Table 9 Data • CMS deploys commands including Create SIM, ECS, FREQ, RA, HCRP, Ad Hoc, and Archive SIM commands • Issuer and HHS receive reports
4	09/11/21	• Data: Final BY20 Data, Proposed BY22 DIY Table 9 Data, and updated transplant, severity, and enrollment duration logic • CMS deploys commands including Create SIM, ECS, FREQ, RA, HCRP, Ad Hoc, and Archive SIM commands • Issuer and HHS receive reports

Similar to the EDGE server reports issuers receive for the applicable (un-simulated) benefit year of risk adjustment, this simulation exercise will also provide issuers with detailed individual-level data and outbound data files for the simulated results,[109] while HHS will only receive summary-level plan data. HHS intends to publish summary information of this transfer simulation exercise, including information on the change in risk score and change in the transfer as a percent of premium at the state market risk pool level.[110] If an issuer wishes to use the simulation results to assist in assessing the impact of the model specification changes on future benefit year transfer amounts, it should do so with caution and in combination with other significant data. In particular, smaller issuers may experience a wider degree of variation, given the impact that larger issuers have on transfers within a state market risk pool.

5.3 Conclusion

CMS constantly considers ways to refine the risk adjustment models and program requirements, including a variety of different options to address prediction concerns for certain subpopulations. This technical paper focuses on ways to improve the current models' predictive accuracy for the lowest-risk enrollees, certain partial-year adult enrollees, and the very highest-risk enrollees. This paper provides additional data and analysis on proposed model specification changes proposed in the proposed 2022 Payment Notice: two-stage weighting, HCC-contingent enrollment duration factors, and interacted HCC counts.

As discussed above, to further assist stakeholders in understanding the impact of these proposed model specification changes, we intend to release the results from the transfer simulation towards the end of 2021 to provide issuers with illustrative information of what transfers would have been with and without the model specification changes discussed in Chapters 2 - 4.

We are accepting comments on this paper (including the Appendices) at RARIPAYMENTOPERATIONS@cms.hhs.gov with the subject line of "2021 Model Update Technical Paper" until November 26, 2021. We also intend to hold a webinar through REGTAP in November 2021 to review this paper.

[109] Issuers need to participate by running the commands provided by CMS within the designated timeframes in order to receive the issuer-level simulation reports.

[110] Note that the transfer estimates from this simulation will not account for transfer impacts due to 2020 benefit year discrepancies.

Appendix A: Cost-Sharing Reduction Induced Demand Factors

This appendix describes the cost-sharing reduction induced demand factors (CSR IDFs)[111] that are currently used in the state payment transfer formula in the HHS risk adjustment methodology, stakeholder feedback regarding the current CSR IDFs, and potential options—that likely would be implemented no earlier than the 2024 benefit year—to improve prediction of CSR enrollees' plan liability in the state payment transfer formula.

Given that no changes to the CSR IDFs were proposed in the 2022 Payment Notice and the changes proposed in that rule are the primary focus of this paper, none of the data or analyses in Sections 1 – 5 of this paper or in the forthcoming EDGE data transfer simulation will reflect any changes to the CSR IDFs. However, we wanted to share information with stakeholders about our ongoing study and consideration of the current CSR IDFs, including potential options we are exploring to update the CSR IDFs. We continue to consider these options and would propose any changes for future benefit years through notice-and-comment rulemaking, likely no earlier than the 2024 benefit year.

A.1 Background on the CSR Induced Demand Factors

Under the ACA, health insurance plans offered in the individual and small group (including merged) markets are generally grouped into metal levels defined by the plans' AV, meaning the generosity of plan benefits, —bronze (60 percent AV), silver (70 percent AV), gold (80 percent AV), and platinum (90 percent AV). Individuals with household income below 250 percent of the federal poverty level (FPL) and enrolled through the Exchanges in silver plan variants are eligible to receive CSR subsidies,[112] which increase their plan's AV beyond 70 percent by providing the enrollee with more generous cost sharing (maximum out of pocket costs, deductibles, copays, etc.) than the standard silver plan. Until October 2017, issuers received payments from the federal government to compensate for the CSRs issuers provided between the standard silver plan AV (70 percent AV) and the CSR plan variant AVs, which are 73 percent AV for enrollees with household incomes above 200 percent and no more than 250 percent FPL, 87 percent AV for enrollees with household incomes above 150 percent and no more than 200 percent FPL, and 94 percent AV for enrollees with household incomes at or above 100 percent and no more than 150 percent FPL. Starting in October 2017, the federal

[111] These values are commonly referred to as the CSR induced demand factors, the CSR induced utilization factors, the CSR adjustment factors, and the CSR multipliers. We will use the phrase "CSR induced demand factors" or "CSR IDFs" throughout this appendix.

[112] American Indians and Alaska Natives enrolled in any metal level also receive CSRs through zero cost sharing and limited cost sharing plan variants, as do Premium Assistance Medicaid Alternative Plan enrollees in 94 percent AV and zero cost sharing silver plan variants.

government ceased payments of CSRs to issuers.[113] Many issuers started raising premiums to cover CSR costs starting in 2018.

When we developed the state payment transfer formula in the 2014 Payment Notice,[114] we determined that we should account for induced demand associated with two separate sources: (1) the plan generosity differences between metal levels and (2) enrollee receipt of CSRs, which result in greater generosity of CSR plan variations relative to standard plans. Both the higher AVs provided by more generous metal levels and the enhanced cost sharing reductions provided by CSR plan variations allow consumers enrolled in these plans to experience lower cost sharing than less generous metal levels or the standard plan with which the CSR plan variations are associated. As such, based on research findings showing that individuals with lower cost sharing consume more medical care and have higher expenditures than individuals with higher cost sharing,[115] we added metal and CSR IDFs to the state payment transfer formula to account for anticipated increased demand associated with enrollees' selection of higher plan metal levels and enrollee receipt of CSRs. We have maintained the same metal and CSR IDFs since the beginning of the program.[116] The state payment transfer formula appears below.[117]

$$T_i = \left[\frac{PLRS_i \cdot metalIDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot metalIDF_i \cdot GCF_i)} - \frac{AV_i \cdot ARF_i \cdot metalIDF_i \cdot GCF_i}{\sum_i (s_i \cdot AV_i \cdot ARF_i \cdot metalIDF_i \cdot GCF_i)}\right] \overline{P_s}$$

And:

$$PLRS_i = \frac{\sum_e (RS_e * csrIDF_e * EMM_e)}{\sum_e BMM_e}$$

Where:

T_i = plan i's per PMPM transfer amount;

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

RS_e = the model calculated risk score for an enrollee's enrollment period e within plan i;

[113] On October 12, 2017, the Attorney General issued a legal opinion that HHS did not have a Congressional appropriation with which to make CSR payments. The Attorney General's opinion regarding CSR payments (2017) is available at: https://www.hhs.gov/sites/default/files/csr-payment-memo.pdf.

[114] 78 FR 15409 at 15421.

[115] See, for example, the RAND Health Insurance Experiment, available at: https://www.rand.org/health-care/projects/hie.html.

[116] This appendix focuses on the CSR IDFs, as HHS has not noted any significant prediction concerns with the metal IDFs and stakeholders' comments have not focused on the metal IDFs.

[117] Previously published versions of the state payment transfer formula, and the version that appears in section 1.2.3 earlier in this paper, do not show separate IDFs for metal level and CSRs. As this appendix focuses specifically on the CSR IDFs to further illustrate the issues under consideration, we have made explicit in this appendix that the IDF term in the previously published formula refers to metal IDFs, and that CSR IDFs are present in the PLRS term. The different display does not modify or otherwise represent a change to the state payment transfer formula as finalized through the Payment Notice rulemaking process.

$csrIDF_e$ = the CSR induced demand factor for an enrollee's enrollment period e within plan i;[118]

EMM_e = the total enrollment months for an enrollee's enrollment period e within plan i;

BMM_e = the billable months[119] for an enrollee's enrollment period e within plan i;

$metalIDF_i$ = plan i's metal induced demand factor;

GCF_i = plan i's geographic cost factor;

s_i = plan i's share of state enrollment;

AV_i = plan i's metal level AV;

ARF_i = allowable rating factor; and

$\bar{P}_S$ = statewide average premium.

As discussed in Section 1.2.3, the first half of the state payment transfer formula ($\frac{PLRS_i \cdot metalIDF_i \cdot GCF_i}{\sum_i (s_i \cdot PLRS_i \cdot metalIDF_i \cdot GCF_i)}$) is referred to as the risk term and includes IDFs that account for both metal and CSR induced demand (CSR IDFs are included in the calculation of the PLRS), so that the risk scores for CSR enrollees are adjusted by both IDFs. For example, a silver CSR enrollee in an 87 or 94 percent AV plan variant would have a 1.03 metal IDF and a 1.12 CSR IDF, as seen in Tables A.1 and A.2 below. Due to the multiplicative nature of the variables in the risk term, the effects of these IDFs combine in the risk term to roughly equal the effect of the 1.15 metal IDF for platinum plans.

The second half of the state payment transfer formula ($\frac{AV_i \cdot ARF_i \cdot metalIDF_i \cdot GCF_i}{\sum_i (s_i \cdot AV_i \cdot ARF_i \cdot metalIDF_i \cdot GCF_i)}$), referred to as the rating term, is an estimate of what an issuer's premium would need to be in the absence of risk selection[120] and therefore only includes plan-level values, including an IDF variable that only accounts for metal induced demand (i.e., the metal IDFs). Note that the rating term does not include a plan's risk score, because rating on health status risk is not allowable under the ACA for individual and small group non-grandfathered health insurance coverage.[121]

---

[118] CSR plan variants are denoted at the 16-digit plan ID level and are subsequently aggregated to the 14-digit plan ID level in the $PLRS_i$ calculation step. Transfers are then calculated using the state payment transfer formula at the 14-digit plan ID level.

[119] When the plan average PLRS is calculated, all plan enrollees are counted in the numerator, but only billable plan enrollees (parents and up to the 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. See https://www.cms.gov/mmrr/Downloads/MMRR2014_004_03_a03.pdf and https://www.cms.gov/mmrr/downloads/mmrr2014_004_03_a04.pdf.

[120] See https://www.cms.gov/mmrr/downloads/mmrr2014_004_03_a04.pdf

[121] See section 2701 of the Public Health Service Act and 45 CFR 147.102.

Because the CSR IDFs are only represented as a component of the plan's risk score, there is no representation of CSR IDFs in the rating term.

Table A.1 shows the CSR IDFs currently used in the state payment transfer formula. Table A.2 shows the metal IDFs currently used in the state payment transfer formula.

Table A.1: CSR Induced Demand Factors

Household Income	Plan AV	Induced Demand 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 A.2 Metal Induced Demand Factors

Metal Level	Induced Demand Factor
Catastrophic	1.00
Bronze	1.00
Silver	1.03
Gold	1.08
Platinum	1.15

A.2 Stakeholder Feedback on CSR Induced Demand Factors

HHS received comments to the proposed 2022 Payment Notice and past Payment Notices asking HHS to review the CSR IDFs in the risk adjustment models. Some of these stakeholders shared anecdotal reports from issuers that suggest the CSR IDFs in the risk term may overcompensate silver plans and undercompensate plans at other metal levels, suggesting this occurs because CSR enrollees tend to be more price sensitive due to their lower incomes and are not in fact using more services despite having more generous coverage.[122] These commenters argued that issuers are setting the silver plan premiums low to attract these CSR enrollees due to a perceived overcompensation in the risk term in the state payment transfer formula that may occur for some issuers. They stated that these silver premiums, which they explained are lower than the commenters' expectations, ultimately reduced the cost of the second lowest cost silver plan that forms the basis of the premium tax credit (PTC) calculation, thereby reducing the amount of PTC that consumers were eligible to receive. The commenters suggested that HHS should take action to address this misalignment, which they contend would reduce premiums costs for 97 percent of Exchange enrollees. Other commenters generally requested that we reevaluate the CSR IDFs to confirm the current values are accurate or to update them based on more recent data, as we have not recalibrated the IDF factors since the inception of the HHS-operated risk adjustment program in the 2014 benefit year.[123] Although we have not updated the CSR IDFs, we have continued to analyze induced demand and CSR plan liability in both the risk term and the rating term of the state payment transfer formula, which we discuss in detail below.

A.2.1 Risk Term Analysis

To consider the impact of the CSR IDFs on the risk term, we tested for induced demand of CSR enrollees using more recently available data. Using the 2016 and 2017 enrollee-level EDGE datasets,[124] we estimated the mean expenditure differentials associated with enrollment in various metal levels and CSR variants to approximate the effect of health insurance coverage on utilization and expenditures. We included the full set of applicable risk adjustment model factors, including demographic factors, HCCs, RXCs (adult models only), and enrollment duration factors (adult models only), in this analysis, to control for those factors that could independently influence enrollee risk and utilization. The results of this analysis using the 2017 individual market enrollee-level EDGE dataset appear in Tables A.3 and A.4 below. In Table A.3, which assesses CSR IDFs, positive values indicate higher expenditures for CSR enrollees compared to non-CSR silver enrollees, while negative values indicate lower expenditures for CSR enrollees. Accordingly, our analysis shows that all CSR silver enrollees except the American Indian/Alaska Native CSR plan variant enrollees demonstrate lower expenditures than non-CSR silver enrollees.[125] This implies a lack of evidence of higher induced demand associated with receipt of CSRs for most CSR enrollees.[126]

---

[122] See Lissenden et al, A Comparison of Health Risk and Costs Across Private Insurance Markets, Med Care 58(2), February 2020, available at: https://pubmed.ncbi.nlm.nih.gov/31688571/.
https://pubmed.ncbi.nlm.nih.gov/31688571/.

[123] See 78 FR 15410 at 15421-2.

[124] Preliminary results using the 2018 and 2019 enrollee-level EDGE datasets suggest consistent results with those presented here.

[125] In addition to the 73 percent, 87 percent, 94 percent, zero cost sharing, and limited cost sharing CSR plan variants, we also tested induced demand for the Premium Assistance Medicaid Alternative Plans (94 percent AV and zero cost sharing variants) and again did not find any evidence of induced demand associated with those enrollees.

[126] We also assessed metal IDFs and found that expenditure differentials indicated that the current metal IDFs are reasonable.

Table A.3: Mean Expenditure Differentials by CSR Plan Variant as Compared to Standard Silver Plan, 2017 EDGE Individual Market Enrollee-Level EDGE Dataset

Household Income	Plan Variant	% Difference in Expenditures
>250% of FPL	On-Exchange Silver Standard Plan (70% AV)	(point of comparison)
<300% of FPL	American Indian/Alaska Native Plan (Zero Cost Sharing)	7.53%
>300% of FPL	American Indian/Alaska Native Plan (Limited Cost Sharing)	3.72%
200-250% of FPL	73% AV Plan Variant	-1.38%
150-200% of FPL	87% AV Plan Variant	-0.74%
100-150% of FPL	94% AV Plan Variant	-1.81%

After testing for induced demand, we tested whether CSR plan variants were being overcompensated by the CSR IDFs as some stakeholders had suggested. However, for this analysis, we had concerns about the limitations of assessing the accuracy of model predictions for CSR enrollees using only simulated silver plan liability. As mentioned earlier in this paper, we calculate PRs as the ratio of predicted to actual weighted mean plan liability expenditures. For most of our analyses, we calculate predicted mean plan liability expenditures (the numerator of the PR) using only the model coefficients to which enrollees are subject. As such, predicted mean plan liability expenditures in this context excludes adjustments for CSR IDFs, metal IDFs, and GCFs for each rating area. On the other hand, we generally simulate actual weighted mean plan liability expenditures (the denominator of the PR) based on standardized benefit designs for each metal, cost trends (e.g., medical inflation) from the years used to calibrate the model to the payment year for which the model is effective, and adjustments to account for the high-cost risk pool. These PRs are an excellent tool for investigating how well the risk adjustment models themselves predict expenditures for various subgroup, but in the state payment transfer formula, predicted plan liability represented by the risk term reflects more than just the risk score from the risk adjustment models. As such, we created an additional type of PR that can be used to evaluate the performance of the relative risk measure (i.e., the risk term) in the state payment transfer formula. We termed this type of PR the "nationally-approximated risk term PR."

Rather than reflecting risk score as derived from the model coefficients exclusively, these PRs are a more complete reflection of the risk term in the state payment transfer formula (i.e. they take into account metal-level IDFs and CSR IDFs, but due to data limitations, do not account for the geographic components of the risk term—specifically, GCF and state-level standardization—because all geographic identifiers were excluded from the 2019 Enrollee-level

EDGE dataset for the individual market).[127] The national weighted average of this more complete reflection of the risk term serves as the numerator in the PR calculation, with the national weighted average of standardized paid claims in the denominator (rather than simulated plan liability expenditures).

The results of the analysis on nationally-approximated risk term PRs, seen in Table A.4, indicate that the most common high CSR variants are predicting actual plan liability relatively accurately on average. Although the 73 percent AV Silver CSR variant is overpredicted by about 14 percent, we note that this CSR variant reflects less than half of the enrollment in the 87 percent and 94 percent AV CSR variants, which demonstrate a fair degree of accuracy as indicated by their PRs.

Table A.4: Nationally-Approximated Risk Term Predictive Ratios by Metal-CSR Level, 2019 Enrollee-level EDGE Individual Market Data and 2019 EDGE Plan-Level Data

Metal-CSR Level	Predictive Ratio	Enrollment Months in Plan Level Data
On-Exchange Silver, no CSR	1.03	10,053,099
Silver CSR, 73% AV	1.14	8,044,442
Silver CSR, 87% AV	1.05	18,752,341
Silver CSR, 94% AV	0.98	30,801,803
American Indian/Alaska Native, Zero Cost Sharing	0.71	478,753
American Indian/Alaska Native, Limited Cost Sharing	0.97	72,995

Although we did not specifically identify any evidence of induced demand associated with CSR enrollees, Table A.4 above shows that the risk term PRs[128] associated with CSR enrollees are mostly adequate (PRs within +/– 5 percent are considered reasonable), at least for the most common high CSR variants, although they could be improved. This suggests that, even if silver plan risk scores from the risk adjustment models are overpredicted relative to the 70 percent AV of the standard silver plan, issuers are also paying more than the expected silver plan liability for CSR enrollees in terms of actual paid claims under these plans. As such, this analysis

[127] Testing the accuracy of predictions by CSR level using full risk term PRs (including the geographic components) would require using 2019 EDGE plan-level data (which can only reflect risk scores from the final 2019 risk adjustment models), rather than 2019 Enrollee-level EDGE dataset for the individual market (which can reflect risk scores resulting from the risk adjustment model changes discussed throughout the paper). Although we conducted analysis on the 2019 risk adjustment model full risk term PRs derived from plan-level data, we have not discussed these results in detail here because they are less relevant in the context of the potential model changes presented in this paper. However, we note that the results were extremely similar to the results in Table A.4.

[128] Risk term PRs, which measure the accuracy of the risk term at *predicting actual paid claims*, are better than risk score PRs for this analysis, since CSR plan variants owe more than 70 percent of allowed charges for their CSR enrollees. As stated in section 1.4, risk score PRs evaluate the accuracy of the PLRS at predicting *simulated PMPM plan liability* based on the silver model factors, which account for 70 percent cost sharing.

suggests the risk term is predicting what plans are actually paying for CSR enrollees reasonably well.

As a result of this analysis, we conclude that, in aggregate, the current CSR IDFs in the risk term are not resulting in meaningful overcompensation of CSR plan variants or under-compensation of plans at other metal levels, contrary to concerns of some stakeholders. Nevertheless, we believe that the current CSR IDFs should be re-evaluated because of: (1) a potential lack of alignment between the original intent of the CSR IDFs (accounting for anticipated induced demand due to lower cost sharing) and their current function; and (2) the desire to refine the CSR IDFs in the risk term using more recent data. See A.3.1 for details on current risk term update options under consideration.

A.2.2 Rating Term Analysis

As previously mentioned, the rating term does not include any adjustment for CSR plan variants because CSR IDFs are assessed as a part of plan liability (the risk term), whereas the rating term reflects the premium that the plan would be allowed to collect from enrollees who are of average risk (taking into account their age, geographic location, the stated AV of their metal level, and the applicable metal IDF). However, the value of the AV in the state payment transfer formula for all silver plans, regardless of whether they are CSR plan variants, is the standard silver plan AV of 70 percent. It therefore does not reflect the higher AV ranges of the CSR silver plan variants, which some issuers are reflecting in higher rates. This AV, which is lower than the liability plans incur for the higher-AV CSR variants, could result in silver plans receiving larger risk adjustment payments or owing smaller charges than they would if the rating term were adjusted for CSR liability applicable to the higher AV CSR plan variants. This lack of a CSR adjustment in the rating term could increase compensation for silver plans and decrease compensation for plans at other metal levels relative to what these plans' transfers would have been in the absence of the higher than 70 percent AV among CSR silver plan variations. Therefore, we intend to consider options to address the current treatment of the CSRs in the rating term along with the re-evaluation of the CSR IDFs in the risk term of the state payment transfer formula. See section A.3.2 for details on current rating term update options under consideration.

A.3 Potential Future Changes to CSR Induced Demand Factors

We have identified several potential options to update the risk term and one option to update the rating term to more precisely account for CSR plan liability in the state payment transfer formula. These options could be done in combination to update both the risk term and rating term together, although these options have varying potential impacts on risk adjustment transfers and the state market risk pools. These options are targeted refinements that reflect the current lack of an appropriation for HHS to make CSR payments to issuers, so several of these options would likely not be appropriate if an appropriation to make CSR payments were to be enacted in the future. As noted above, we continue to consider these options and would propose any changes for future benefit years through notice-and-comment rulemaking, likely no earlier than the 2024 benefit year.

A.3.1 Risk Term Update Options

The following three options would revise the CSR IDFs in the risk term[129] to more precisely account for CSR plan liability.

1. Reframe and Recalibrate the CSR IDFs: Under this option, we would reframe and recalibrate the current CSR IDFs to shift their focus from accounting for anticipated induced demand to accounting for the higher AV of CSR plan variants. This change would also update the CSR IDFs using enrollee-level EDGE data for the first time, as they were developed and calibrated using MarketScan® large group market data and have not been updated since 2015. Specifically, under this option, we would refer to the CSR IDFs as "CSR adjustment factors" and refine their values by recalibrating them for all CSR plan variants such that the risk term PRs shown in Table A.4 are as close as possible to 1.00 (accurate prediction) for all CSR plan variants.[130] We believe that this option is consistent with our recent analysis suggesting that the anticipated induced demand does not exist for these enrollees and that this option would address stakeholder requests to evaluate the current CSR IDFs. This option would also preserve flexibility to update the risk term for CSR enrollees in the future. However, as the potential new recalibrated CSR Adjustment Factors could be higher than the current CSR IDFs, this option might not have the impact on transfers that some stakeholders have argued is needed. A potential set of recalibrated CSR adjustment factors appears in Table A.5 below.[131]

[129] These options do not include any changes to the metal IDFs as our analysis found the current metal IDFs in the risk term are reasonable.
[130] Under this option, we would calculate the CSR adjustment factors using multiple years of enrollee-level EDGE data to ensure consistency but would not recalibrate their values each year.
[131] The potential set of recalibrated CSR adjustment factors were recalibrated using only 2019 EDGE plan-level data and are being provided for illustrative purposes. As part of our further consideration of this option, HHS intends to analyze additional years of EDGE plan-level data to ensure consistency of factors across data years.

Table A.5: Potential Recalibrated CSR Adjustment Factors

Household Income	Plan AV	Adjustment Factor
Silver Plan Variant Recipients		
100-150% of Federal Poverty Line (FPL)	Plan Variation 94%	1.20
150-200% of FPL	Plan Variation 87%	1.10
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%)	1.15
<300% of FPL	Gold (80%)	1.35
<300% of FPL	Silver (70%)	1.45
<300% of FPL	Bronze (60%)	1.55
Limited Cost Sharing Recipients		
>300% of FPL	Platinum (90%)	1.00
>300% of FPL	Gold (80%)	1.05
>300% of FPL	Silver (70%)	1.15
>300% of FPL	Bronze (60%)	1.25

2. Use Platinum Risk Adjustment Model Factors: The second option would eliminate the CSR IDFs from the PLRS component of the risk term and use the platinum risk adjustment model factors to calculate risk scores for CSR enrollees in the 87 percent, 94 percent, zero cost sharing, and limited cost sharing plan variants. Although we would be calculating PLRS for these enrollees based on the platinum model factors, the metal IDFs would remain unchanged and continue to reflect the metal level of the CSR enrollees (i.e. the silver metal IDFs for the 87 percent and 94 percent AV CSR variants). This option would better capture plan liability for those enrollees since the platinum AV (90 percent) is much closer to the true AV of those plan variants. Under this option, we would continue to use the silver risk adjustment model factors to calculate risk scores for CSR enrollees in the 73 percent AV plan variants. Some stakeholders have requested this option, and we have considered it in the past.[132] However, it is less flexible than the previous option in terms of capturing differences among CSR plan variant enrollees (for example, this would not account for differences between 87 percent and 94 percent CSR plan variant enrollees, or between those enrollees and the zero-cost sharing and limited cost sharing CSR plan variant enrollees).
3. Create Separate Risk Adjustment Models for CSR Plan Variants: The third option would eliminate the CSR IDFs from the PLRS component of the risk term and create separate risk adjustment models for the 87 percent, 94 percent, zero cost sharing, and limited cost sharing CSR plan variants. These additional risk adjustment models

[132] See the August 10, 2017 FAQ titled Information on Risk Adjustment Methodology and Rate Filing Deadlines, available at: https://www.regtap.info/uploads/library/QHP_FAQ_5CR_081117.pdf.

would account for the difference in plan liability between enrollees in these CSR plan variants and the standard silver plan and 73 percent CSR plan variant enrollees, and it would result in risk scores that are more aligned with actual plan liability for CSR enrollees. Some stakeholders have also requested this approach in the past. However, it would complicate risk adjustment by adding several new models to the HHS-operated methodology. Moreover, the HHS risk adjustment models are calibrated using the same standard population across all models. This standard population is drawn from enrollee-level EDGE data (including data from the small group market) and varies by model only in the set of standard payment parameters used to simulate plan liability by metal level. However, it is difficult to justify either (1) using the standard population inclusive of small group market data to calibrate CSR plans because CSR plans are only offered on the individual market, or (2) using a CSR- or individual-market-specific population when all other models are based on the standard population. We would also have to determine if we should have new models for each age group (adult, child, and infant) for each of the 87 percent, 94 percent, zero cost sharing, and limited cost sharing CSR plan variants, or if we should consolidate these models (and if so, how) to increase model factor stability. Furthermore, these higher-AV CSR plan variants have overlapping *de minimis* AV ranges with platinum plans, so the new model factors would be very similar to the existing platinum model factors. Therefore, we believe it would be simpler to adopt one of the other risk term update options described in the previous paragraphs.

A.3.2 Rating Term Update Options

To more precisely reflect CSR plan variants in the second half of the state payment transfer formula, we could use the average silver plan AV of an issuer's silver plan, including CSR plan variant AVs, weighted by the proportion of the issuer's plan enrollment that falls under each CSR plan variant, in the rating term rather than use the standard silver AV (70 percent), as is currently the case. To determine a single rating term AV for a silver plan, we could calculate an enrollment-weighted average AV across the five main silver plan variants.[133] However, if the weighted average were dependent on plan enrollment for each variant and varied at the 14-digit plan ID level, the plan's rating term AV (and thus expected premium) would vary based on plan selection by CSR enrollees, which is contrary to the purpose of risk adjustment. Therefore, to avoid this concern, we could calculate the weighted average rating term AV across all plans at the national level, thereby ensuring the AV will be the same amount for all silver plans nationally. An approach that reflects the higher AV for the CSR plan variants in both terms

[133] The AVs of the five main silver plan variants are the 70 percent AV for the off-Exchange silver plan, the 70 percent AV for the on-Exchange standard silver plan, 73 percent AV for the CSR plan variant for individuals with household income 200-250 percent FPL, 87 percent AV for the CSR plan variant for individuals with household income 150-200 percent FPL, and 94 percent AV for the CSR plan variant for individuals with household income 100-150 percent FPL.

would be in keeping with the intent of the AV factor in the state payment transfer formula to reflect differences in plan generosity rather than selection.

A.4 Next Steps

Our intention with this appendix is to familiarize stakeholders with our considerations regarding the current CSR and metal IDFs and potential options we are considering to more precisely predict plan liability for CSR enrollees. As noted above, we did not propose changes to the CSR or metal IDFs in the 2022 Payment Notice and these options will not be reflected in the EDGE server transfer simulation HHS intends to provide later this year. We continue to conduct analyses of these options and will propose any changes in future notice-and-comment rulemaking, likely no earlier than the 2024 benefit year. As noted above, we are also accepting comments on this paper (including the Appendices) at RARIPAYMENTOPERATIONS@cms.hhs.gov with the subject line of "2021 Model Update Technical Paper" until November 26, 2021.

Appendix B: List of Acute HCCs Used in Enrollment Duration Factors Analysis

The below list includes the HCCs we considered to be acute for the purposes of the analysis of enrollment duration factors by presence of acute vs. chronic HCCs found in section 3.3.3.[134]

(1) HCC 2: Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock
(2) HCC 3: Central Nervous System Infections, Except Viral Meningitis
(3) HCC 4: Viral or Unspecified Meningitis
(4) HCC 23: Protein-Calorie Malnutrition
(5) HCC 38: Acute Liver Failure/Disease, Including Neonatal Hepatitis
(6) HCC 42: Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis
(7) HCC 45: Intestinal Obstruction
(8) HCC 47: Acute Pancreatitis/Other Pancreatic Disorders and Intestinal Malabsorption
(9) HCC 54: Necrotizing Fasciitis
(10) HCC 55: Bone/Joint/Muscle Infections/Necrosis
(11) HCC 63: Cleft Lip/Cleft Palate
(12) HCC 126: Respiratory Arrest
(13) HCC 127: Cardio-Respiratory Failure and Shock, Incl. Respiratory Distress Syndromes
(14) HCC 135: Heart Infection/Inflammation, Except Rheumatic
(15) HCC 154: Vascular Disease with Complications
(16) HCC 156: Pulmonary Embolism and Deep Vein Thrombosis
(17) HCC 163: Aspiration, Specified Bacterial Pneumonias and Other Severe Lung Infections
(18) HCC 203: Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism
(19) HCC 204: Miscarriage with Complications
(20) HCC 205: Miscarriage with No or Minor Complications
(21) HCC 207: Completed Pregnancy With Major Complications
(22) HCC 208: Completed Pregnancy With Complications
(23) HCC 209: Completed Pregnancy with No or Minor Complications
(24) HCC 226: Hip Fractures and Pathological Vertebral or Humerus Fractures
(25) HCC 227: Pathological Fractures, Except of Vertebrae, Hip, or Humerus

[134] As this analysis was conducted using the 2017 enrollee-level EDGE dataset, this list of HCCs comes from the V05 classification model and does not represent the current (2021 benefit year) list of payment HCCs in the V07 classification model.

Appendix C: HCCs in the 2021 Adult Model Severity Illness Indicator Variable and HCCs Selected for HCC Interacted Counts Variables for the Adult and Child Models

Payment HCC		2021 Model Severity Illness Indicator	New Severe Illness Indicator	New Transplant Indicator
HCC 2	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	X	X	
HCC 3	Central Nervous System Infections, Except Viral Meningitis		X	
HCC 4	Viral or Unspecified Meningitis		X	
HCC 6	Opportunistic Infections		X	
HCC 18	Pancreas Transplant Status[135]		X	X
HCC 23	Protein-Calorie Malnutrition		X	
HCC 34	Liver Transplant Status/Complications		X	X
HCC 41	Intestine Transplant Status/Complications		X	X
HCC 42	Peritonitis/Gastrointestinal Perforation/Necrotizing Enterocolitis	X	X	
HCC 96	Prader-Willi, Patau, Edwards, and Autosomal Deletion Syndromes		X	
HCC 120	Seizure Disorders and Convulsions	X		
HCC 121	Hydrocephalus		X	
HCC 122	Coma, Brain Compression/Anoxic Damage	X	X	
HCC 125	Respirator Dependence/Tracheostomy Status	X	X	
HCC 126	Respiratory Arrest	X		

[135] HCC 18 (Pancreas Transplant) was not included in this list in the proposed 2022 Payment Notice. HCC 18 has a much lower coefficient than any of the other transplant HCCs in the adult models and was not underpredicted by the models. In the proposed 2022 Payment Notice, we proposed to exclude it from the list in and solicited comments on the proposed treatment of HCC 18. We did not receive any comments on including or excluding HCC 18 in the severity and transplant indicators. We included it in the analysis is in this paper to be consistent with all the other transplant HCCs.

Payment HCC	2021 Model Severity Illness Indicator	New Severe Illness Indicator	New Transplant Indicator	Payment HCC
HCC 127	Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes	X		
HCC 135	Heart Infection/Inflammation, Except Rheumatic		X	
HCC 145	Intracranial Hemorrhage		X	
HCC 156	Pulmonary Embolism and Deep Vein Thrombosis	X	X	
HCC 158	Lung Transplant Status/Complications		X	X
HCC 163	Aspiration and Specified Bacterial Pneumonias and Other Severe Lung Infections		X	
HCC 183	Kidney Transplant Status/Complications		X	X
HCC 218	Extensive Third Degree Burns		X	
HCC 223	Severe Head Injury		X	
HCC 251	Stem Cell, Including Bone Marrow, Transplant Status/Complications		X	X
G13	(Includes HCC 126 Respiratory Arrest and HCC 127 Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes)		X	
G14	(Includes HCC 128 Heart Assistive Device/Artificial Heart and HCC 129 Heart Transplant Status/ Complications)		X	X