[Billing Code: 4120-01-P]

**DEPARTMENT OF HEALTH AND HUMAN SERVICES**

**45 CFR Part 153**

**[CMS-9913-P]**

**RIN 0938-AU23**

# Amendments to the HHS-operated Risk Adjustment Data Validation under the Patient Protection and Affordable Care Act's HHS-operated Risk Adjustment Program

**AGENCY:** Centers for Medicare & Medicaid Services (CMS), Department of Health and Human Services (HHS).

**ACTION:** Proposed rule.

**SUMMARY:** This rule proposes to adopt certain changes to the risk adjustment data validation error estimation methodology starting with the 2019 benefit year and beyond for states where the Department of Health and Human Services (HHS) operates the risk adjustment program. The Patient Protection and Affordable Care Act (PPACA) established a permanent risk adjustment program under which payments are made to health insurance issuers that attract higher-than-average risk populations funded by payments from health insurance issuers that attract lower-than-average risk populations. To ensure the integrity of the HHS-operated risk adjustment program, CMS, on behalf of HHS, performs risk adjustment data validation, also known as HHS-RADV, to validate the accuracy of data submitted by issuers for the purposes of risk adjustment transfer calculations. Based on lessons learned from the first payment year of HHS-RADV, this rule proposes changes to the HHS-RADV error estimation methodology, which is used to calculate adjusted risk scores and risk adjustment transfers, beginning with the 2019 benefit year of HHS-RADV. This rule also proposes to change the benefit year to which HHS-RADV adjustments to risk scores and risk adjustment transfers would be applied starting

with 2021 benefit year HHS-RADV. These proposals seek to further the integrity of the HHS-RADV program, address stakeholder feedback, promote fairness, and improve the predictability of HHS-RADV adjustments.

**DATES:** To be assured consideration, comments must be received at one of the addresses provided below, no later than 5 p.m. on **[**Insert date 30 days after date of publication in the **Federal Register]**.

**ADDRESSES:** In commenting, please refer to file code CMS-9913-P. Because of staff and resource limitations, we cannot accept comments by facsimile (FAX) transmission.

Comments, including mass comment submissions, must be submitted in one of the following three ways (please choose only one of the ways listed):

1. Electronically. You may submit electronic comments on this regulation to **http://www.regulations.gov**. Follow the "Submit a comment" instructions.

2. By regular mail. You may mail written comments to the following address ONLY:

   Centers for Medicare & Medicaid Services,

   Department of Health and Human Services,

   Attention: CMS-9913-P,

   P.O. Box 8010,

   Baltimore, MD 21244-8010.

Please allow sufficient time for mailed comments to be received before the close of the comment period.

3. By express or overnight mail. You may send written comments to the following address ONLY:

   Centers for Medicare & Medicaid Services,

Department of Health and Human Services,

Attention: CMS-9913-P,

Mail Stop C4-26-05,

7500 Security Boulevard,

Baltimore, MD 21244-1850.

For information on viewing public comments, see the beginning of the "SUPPLEMENTARY INFORMATION" section.

**FOR FURTHER INFORMATION CONTACT:** Allison Yadsko, (410) 786-1740; Joshua Paul, (301) 492-4347; Adrianne Patterson, (410) 786-0686; and Jaya Ghildiyal, (301) 492-5149.

**SUPPLEMENTARY INFORMATION:**

Inspection of Public Comments: All comments received before the close of the comment period are available for viewing by the public, including any personally identifiable or confidential business information that is included in a comment. We post all comments received before the close of the comment period on the following Web site as soon as possible after they have been received: http://www.regulations.gov. Follow the search instructions on that Web site to view public comments.

## I. Background

### A. Legislative and Regulatory Overview

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) was

enacted on March 30, 2010. These statutes are collectively referred to as "PPACA" in this proposed rule. Section 1343 of the PPACA[1] established a permanent risk adjustment program to provide payments to health insurance issuers that attract higher-than-average risk populations, such as those with chronic conditions, funded by payments from those that attract lower-than-average risk populations, thereby reducing incentives for issuers to avoid higher-risk enrollees. The PPACA directs the Secretary, in consultation with the states, to establish criteria and methods to be used in carrying out risk adjustment activities, such as determining the actuarial risk of enrollees in risk adjustment covered plans within a state market risk pool.[2] The statute also provides that the Secretary may utilize criteria and methods similar to the ones utilized under Medicare Parts C or D.[3] Consistent with section 1321(c)(1) of the PPACA, the Secretary is responsible for operating the risk adjustment program on behalf of any state that elected not to do so. For the 2014 – 2016 benefit years, all states and the District of Columbia, except Massachusetts, participated in the HHS-operated risk adjustment program. Since the 2017 benefit year, all states and the District of Columbia have participated in the HHS-operated risk adjustment program.

Data submission requirements for the HHS-operated risk adjustment program are set forth at 45 CFR 153.700 through 153.740. Each issuer is required to establish and maintain an External Data Gathering Environment (EDGE) server on which the issuer submits masked enrollee demographics, claims, and encounter diagnosis-level data in a format specified by HHS.

[1] 42 U.S.C. 18063.
[2] 42 U.S.C. 18063(a) and (b).
[3] 42 U.S.C. 18063(b).

Issuers must also execute software provided by HHS on their respective EDGE servers to generate summary reports, which HHS uses to calculate the enrollee-level risk score to determine the average plan liability risk scores for each state market risk pool, the individual issuers' plan liability risk scores, and the transfer amounts by state market risk pool for the applicable benefit year.

Pursuant to 45 CFR 153.350, HHS performs risk adjustment data validation (also known as HHS-RADV) to validate the accuracy of data submitted by issuers for the purposes of risk adjustment transfer calculations for states where HHS operates the risk adjustment program. This process establishes uniform audit standards to ensure that actuarial risk is accurately and consistently measured, thereby strengthening the integrity of the risk adjustment program.[4] HHS-RADV also ensures that issuers' actual actuarial risk is reflected in risk adjustment transfers and that the HHS-operated program assesses charges to issuers with plans with lower-than-average actuarial risk while making payments to issuers with plans with higher-than-average actuarial risk. Pursuant to 45 CFR 153.350(a), HHS, in states where it operates the program, must ensure proper validation of a statistically valid sample of risk adjustment data from each issuer that offers at least one risk adjustment covered plan[5] in that state. Under 45 CFR 153.350, HHS, in states where it operates the program, may adjust the plan average actuarial risk for a risk adjustment covered plan based on discrepancies discovered as a

[4] HHS also has general authority to audit issuers of risk adjustment covered plans pursuant to 45 CFR 153.620(c).
[5] See 45 CFR 153.20 for the definition of "risk adjustment covered plan."

result of HHS-RADV and use those adjusted risk scores to modify charges and payments to all risk adjustment covered plan issuers in the same state market risk pool.

For the HHS-operated risk adjustment program, 45 CFR 153.630 requires an issuer of a risk adjustment covered plan to have an initial and second validation audit performed on its risk adjustment data for the applicable benefit year. Each issuer must engage one or more independent auditors to perform the initial validation audit of a sample of risk adjustment data selected by HHS.[6] After the initial validation audit entity has validated the HHS-selected sample, a subsample is validated in a second validation audit.[7] The second validation audit is conducted by an entity HHS retains to verify the accuracy of the findings of the initial validation audits.

HHS conducted two pilot years of HHS-RADV for the 2015 and 2016 benefit years[8] to give HHS and issuers experience with HHS-RADV prior to applying HHS-RADV findings to adjust issuers' risk scores, as well as the risk adjustment transfers in the applicable state market risk pool(s). The 2017 benefit year HHS-RADV was the first non-pilot year that resulted in adjustments to issuers' risk scores and the risk adjustment transfers in the applicable state market risk pool(s) as a result of HHS-RADV findings.[9,10]

---

[6] 45 CFR 153.630(b).
[7] 45 CFR 153.630(c).
[8] HHS-RADV was not conducted for the 2014 benefit year. See FAQ ID 11290a (March 7, 2016), available at: https://www.regtap.info/faq_viewu.php?id=11290.
[9] The Summary Report of 2017 Benefit Year HHS-RADV Adjustments to Risk Adjustment Transfers released on August 1, 2019 is available at: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/BY2017-HHSRADV-Adjustments-to-RA-Transfers-Summary-Report.pdf.

When initially developing the HHS-RADV process, HHS sought the input of issuers, consumer advocates, providers, and other stakeholders, and issued the "Affordable Care Act HHS-Operated Risk Adjustment Data Validation Process White Paper" on June 22, 2013 (the 2013 RADV White Paper).[11] The 2013 RADV White Paper discussed and sought comment on a number of potential considerations for the development and operation of the HHS-RADV program. Based on the feedback received, HHS promulgated regulations to implement HHS-RADV that we have modified in certain respects based on experience and public comments, as follows.

In the July 15, 2011 **Federal Register** (76 FR 41929), we published a proposed rule outlining the framework for the risk adjustment program, including standards related to HHS-RADV. We implemented the risk adjustment program and adopted standards related to HHS-RADV in a final rule, published in the March 23, 2012 **Federal Register** (77 FR 17219) (Premium Stabilization Rule). The HHS-RADV regulations adopted in the Premium Stabilization Rule provide for adjustments to risk scores and risk adjustment transfers to reflect HHS-RADV errors, including the two-sided nature of such adjustments.

---

[10] The one exception is for Massachusetts issuers, who were not able to participate in prior HHS-RADV pilot years because the state operated risk adjustment for the 2014-2016 benefit years. Therefore, HHS made the 2017 benefit year HHS-RADV a pilot year for Massachusetts issuers. See 84 FR 17454 at 17508.

[11] A copy of the Affordable Care Act HHS-Operated Risk Adjustment Data Validation Process White Paper (June 22, 2013) is available at: https://www.regtap.info/uploads/library/ACA_HHS_OperatedRADVWhitePaper_062213_5CR_050718.pdf.

In the December 7, 2012 **Federal Register** (77 FR 73117), we published a proposed rule outlining benefit and payment parameters related to the risk adjustment program, including six steps for error estimation for HHS-RADV in 45 CFR 153.630 (proposed 2014 Payment Notice). We published the 2014 Payment Notice final rule in the March 11, 2013 **Federal Register** (78 FR 15436). In addition to finalizing 45 CFR 153.630, this final rule further clarified HHS-RADV policies, including that adjustments would occur when an issuer under-reported its risk scores.

In the December 2, 2013 **Federal Register** (78 FR 72321), we published a proposed rule outlining the benefit and payment parameters related to the risk adjustment program (proposed 2015 Payment Notice). This rule also included several HHS-RADV proposals. We published the 2015 Payment Notice final rule, which finalized HHS-RADV requirements related to sampling; initial validation audit standards, second validation audit processes, and medical record review as the basis of enrollee risk score validation; the error estimation process and original methodology; and HHS-RADV appeals, oversight, and data security standards in the March 11, 2014 **Federal Register** (79 FR 13743). Under the original methodology adopted in that final rule, almost every failure to validate an Hierarchical Condition Category (HCC) during HHS-RADV would have resulted in an adjustment to the issuer's risk score and an accompanying adjustment to all transfers in the applicable state market risk pool.

In the September 6, 2016 **Federal Register** (81 FR 61455), we published a proposed rule outlining benefit and payment parameters related to the risk adjustment program (proposed 2018 Payment Notice) that included proposals related to HHS-RADV. We published the 2018 Payment Notice final rule in the December 22, 2016 **Federal Register** (81 FR 94058), which included finalizing proposals related to HHS-RADV discrepancy reporting, clarifications related

to certain aspects of the HHS-RADV appeals process, and a materiality threshold for HHS-RADV to ease the burden of the annual audit requirements for smaller issuers. Under the materiality threshold, issuers with total annual premiums at or below $15 million are not subject to annual initial validation audit requirements, but would be subject to such audits approximately every 3 years (barring risk-based triggers that would warrant more frequent audits).

In the November 2, 2017 **Federal Register** (82 FR 51042), we published a proposed rule outlining benefit and payment parameters related to the risk adjustment program (proposed 2019 Payment Notice) that included proposed provisions related to HHS-RADV. We published the 2019 Payment Notice final rule in the April 17, 2018 **Federal Register** (83 FR 16930), which included finalizing for 2017 benefit year HHS-RADV and beyond, an amended error estimation methodology to only calculate and adjust issuers' risk scores when an issuer's failure rate is statistically significantly different from other issuers based on three HCC groupings (low, medium, and high), that is, when an issuer is identified as an outlier. We also finalized an exemption for issuers with 500 or fewer billable member months from HHS-RADV; a requirement that initial validation audit samples only include enrollees from state market risk pools with more than one issuer; clarifications regarding civil money penalties for non-compliance with HHS-RADV; and a process to handle demographic or enrollment errors discovered during HHS-RADV. We finalized an exception to the prospective application of

HHS-RADV results for exiting issuers,[12] such that exiting outlier issuers' results are used to adjust the benefit year being audited (rather than the following transfer year).

In the July 30, 2018 **Federal Register** (83 FR 36456), we published a final rule that adopted the 2017 benefit year HHS-operated risk adjustment methodology set forth in the final rules published in the March 23, 2012 and March 8, 2016 editions of the **Federal Register** (77 FR 17220 through 17252 and 81 FR 12204 through 12352, respectively). This final rule set forth additional explanation of the rationale supporting use of statewide average premium in the HHS-operated risk adjustment state payment transfer formula for the 2017 benefit year, including why the program is operated in a budget-neutral manner. This final rule permitted HHS to resume 2017 benefit year program operations, including collection of risk adjustment charges and distribution of risk adjustment payments. HHS also provided guidance as to the operation of the HHS-operated risk adjustment program for the 2017 benefit year in light of publication of this final rule.[13]

In the August 10, 2018 **Federal Register** (83 FR 39644), we published a proposed rule concerning the adoption of the 2018 benefit year HHS-operated risk adjustment methodology set forth in the final rules published in the March 23, 2012 and December 22, 2016 editions of the

[12] To be an exiting issuer, the issuer has to exit all of the market risk pools in the state (that is, not sell or offer any new plans in the state). If an issuer only exits some market risk pools in the state, but continues to sell or offer plans in others, it is not an exiting issuer. A small group issuer with off-calendar year coverage, who exits the small group market risk pool in a state and only has small group carry-over coverage that ends in the next benefit year, and is not otherwise selling or offering new plans in any market risk pools in the state, would be an exiting issuer. See 83 FR 16965 through 16966 and 84 FR 17503. The exiting issuer exception is discussed in Section II.B.

[13] "Update on the HHS-operated Risk Adjustment Program for the 2017 Benefit Year." July 27, 2018. Available at https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/2017-RA-Final-Rule-Resumption-RAOps.pdf .

**Federal Register** (77 FR 17220 through 17252 and 81 FR 94058 through 94183, respectively). The proposed rule set forth additional explanation of the rationale supporting use of statewide average premium in the HHS-operated risk adjustment state payment transfer formula for the 2018 benefit year, including why the program is operated in a budget-neutral manner. In the December 10, 2018 **Federal Register** (83 FR 63419), we issued a final rule adopting the 2018 benefit year HHS-operated risk adjustment methodology as established in the final rules published in the March 23, 2012 and the December 22, 2016 (77 FR 17220 through 1752 and 81 FR 94058 through 94183, respectively) editions of the **Federal Register.** This final rule permitted HHS to resume 2018 benefit year program operations, including collection of risk adjustment charges and distribution of risk adjustment payments.

In the January 24, 2019 **Federal Register** (84 FR 227), we published a proposed rule outlining the benefit and payment parameters related to the risk adjustment program, including updates to HHS-RADV requirements (proposed 2020 Payment Notice). We published the 2020 Payment Notice final rule in the April 25, 2019 **Federal Register** (84 FR 17454). The final rule included policies related to incorporating risk adjustment prescription drug categories (RXCs)[14] into HHS-RADV beginning with the 2018 benefit year and extending the Neyman allocation to the 10$^{th}$ stratum for HHS-RADV sampling. We also finalized using precision analysis to determine whether the second validation audit results of the full sample or the subsample (of up to 100 enrollees) results should be used in place of initial validation audit results when an

[14] An RXC uses a drug to impute a diagnosis (or indicate the severity of diagnosis) otherwise indicated through medical coding in a hybrid diagnoses-and-drugs risk adjustment model.

issuer's initial validation audit results have insufficient agreement with SVA results following a pairwise means test. We clarified the application and distribution of default data validation charges under 45 CFR 153.630(b)(10) and how CMS will apply error rates for exiting issuers and sole issuer markets. We codified the previously established materiality threshold and exemption for issuers with 500 or fewer billable member months and established a new exemption from HHS-RADV for issuers in liquidation who met certain conditions. In response to comments, in the final rule, we updated the timeline for collection, distribution, and reporting of HHS-RADV adjustments to transfers; provided that the 2017 benefit year would be a pilot year for HHS-RADV for Massachusetts; and established that the 2018 benefit year would be a pilot year for incorporating RXCs into HHS-RADV.

In the February 6, 2020 **Federal Register** (85 FR 7088), we published a proposed rule outlining the benefit and payment parameters related to the risk adjustment program (proposed 2021 Payment Notice), including several HHS-RADV proposals. Among other things, in this rulemaking, we proposed updates to the diagnostic classifications and risk factors in the HHS risk adjustment models beginning with the 2021 benefit year to reflect more recent claims data, as well as proposed amendments to the outlier identification process for HHS-RADV in cases where an issuer's HCC count is low. We proposed that beginning with 2019 benefit year HHS-RADV, any issuer with fewer than 30 HCCs (diagnostic conditions) within an HCC failure rate group would not be determined an outlier. We also proposed to make 2019 benefit year HHS-RADV another pilot year for the incorporation of RXCs to allow additional time for HHS, issuers, and auditors to gain experience with validating RXCs. On May 14, 2020, we published the HHS Notice of Benefit and Payment Parameters for 2021 final rule (85 FR 29164) (2021 Payment Notice) that finalized these HHS-RADV changes as proposed. The proposed updates to

the diagnostic classifications and risk factors in the HHS risk adjustment models were also finalized with some modifications.

As explained in prior notice-and-comment rulemaking,[15] while the PPACA did not include an explicit requirement that the risk adjustment program operate in a budget-neutral manner, HHS is constrained by appropriations law to devise and implement its risk adjustment program in a budget-neutral fashion.[16] Although the statutory provisions for many other PPACA programs appropriated funding, authorized amounts to be appropriated, or provided budget authority in advance of appropriations,[17] the PPACA neither authorized nor appropriated additional funding for risk adjustment payments beyond the amount of charges paid in, and did not authorize HHS to obligate itself for risk adjustment payments in excess of charges collected.[18] Indeed, unlike the Medicare Part D statute, which expressly authorized the appropriation of funds and provided budget authority in advance of appropriations to make Part D risk-adjusted payments, the PPACA's risk adjustment statute made no reference to additional appropriations.[19] Congress did not give HHS discretion to implement a risk adjustment program

[15] See, e.g., 78 FR 15441 and 83 FR 16930.
[16] Also see *New Mexico Health Connections v. United States Department of Health and Human Services*, 946 F.3d 1138 (10th Cir. 2019).
[17] For examples of PPACA provisions appropriating funds, see PPACA secs. 1101(g)(1), 1311(a)(1), 1322(g), and 1323(c). For examples of PPACA provisions authorizing the appropriation of funds, see PPACA secs. 1002, 2705(f), 2706(e), 3013(c), 3015, 3504(b), 3505(a)(5), 3505(b), 3506, 3509(a)(1), 3509(b), 3509(e), 3509(f), 3509(g), 3511, 4003(a), 4003(b), 4004(j), 4101(b), 4102(a), 4102(c), 4102(d)(1)(C), 4102(d)(4), 4201(f), 4202(a)(5), 4204(b), 4206, 4302(a), 4304, 4305(a), 4305(c), 5101(h), 5102(e), 5103(a)(3), 5203, 5204, 5206(b), 5207, 5208(b), 5210, 5301, 5302, 5303, 5304, 5305(a), 5306(a), 5307(a), and 5309(b).
[18] See 42 U.S.C. 18063.
[19] Compare 42 U.S.C. 18063 (failing to specify source of funding other than risk adjustment charges), with 42 U.S.C. 1395w-116(c)(3) (authorizing appropriations for Medicare Part D risk adjusted payments); 42 U.S.C. 1395w-115(a) (establishing "budget authority in advance of appropriations Acts" for Medicare Part D risk adjusted payments).

that was not budget neutral. Because Congress omitted from the PPACA any provision appropriating independent funding or creating budget authority in advance of an appropriation for the risk adjustment program, we explained that HHS could not – absent another source of appropriations – have designed the program in a way that required payments in excess of collections consistent with binding appropriations law.

B. Stakeholder Consultation and Input

HHS has consulted with stakeholders on policies related to the HHS-operated risk adjustment program and HHS-RADV. We held a series of stakeholder listening sessions to gather input, and received input from numerous interested groups, including states, health insurance issuers, and trade groups. We also issued a white paper for public comment on December 6, 2019 entitled the HHS Risk Adjustment Data Validation (HHS-RADV) White Paper (2019 RADV White Paper).[20] We considered comments received on the 2019 RADV White Paper and in connection with previous rules as we developed the policies in this proposed rule.

## II. Provisions of the Proposed Regulations

HHS conducts HHS-RADV under 45 CFR 153.630 and 153.350 in any state where HHS is operating risk adjustment on a state's behalf. Since the 2017 benefit year, HHS has been operating risk adjustment and HHS-RADV in all 50 states and the District of Columbia. The purpose of HHS-RADV is to ensure issuers are providing accurate and complete risk adjustment

[20] The 2019 RADV White Paper is available at: https://www.cms.gov/files/document/2019-hhs-risk-adjustment-data-validation-hhs-radv-white-paper.

data to HHS, which is crucial to the purpose and proper functioning of the HHS-operated risk adjustment program. HHS-RADV ensures that issuers' actual actuarial risk is reflected in risk adjustment transfers and that the HHS-operated risk adjustment program assesses charges to issuers with plans with lower-than-average actuarial risk while making payments to issuers with plans with higher-than-average actuarial risk.

HHS-RADV consists of an initial validation audit and a second validation audit. Under 45 CFR 153.630, each issuer of a risk adjustment covered plan must engage an independent initial validation auditor. The issuer provides demographic, enrollment, claims data and medical record documentation for a sample of enrollees selected by HHS to its initial validation auditor for data validation. Each issuer's initial validation audit is followed by a second validation audit, which is conducted by an entity that HHS retains to verify the accuracy of the findings of the initial validation audit.

This rule proposes changes to two aspects of HHS-RADV: (A) the error rate calculation, and (B) the application of HHS-RADV results. Beginning with the 2019 benefit year of HHS-RADV,[21] we propose to: (1) modify the HCC grouping methodology used in the error rate calculation; (2) refine the error rate calculation in cases where an outlier issuer is only slightly outside of the confidence interval for one or more HCC groups; and (3) modify the error rate calculation in cases where a negative error rate outlier issuer also has a negative failure rate. We

[21] As part of the Administration's efforts to combat the Coronavirus Disease 2019 (COVID-19), we announced the postponement of the 2019 benefit year RADV process. We intend to provide further guidance by August 2020 on our plans to begin 2019 benefit year RADV in calendar year 2021. See https://www.cms.gov/files/document/2019-HHS-RADV-Postponement-Memo.pdf.

also propose, beginning with the 2021 benefit year of HHS-RADV, to transition from the current prospective application of HHS-RADV results[22] to an approach that would apply HHS-RADV results to the benefit year being audited. We believe these proposals specifically address stakeholder feedback received after the first payment year of HHS-RADV. These proposals seek to further the integrity of the HHS-RADV program, while promoting fairness and improving the predictability of HHS-RADV.

In addition to soliciting comments on the following proposals, we also request feedback on the potential impact of the COVID-19 public health emergency on the proposed timelines for implementation of the proposals in this rulemaking.

A. Error Rate Calculation Methodology

HHS recognizes that variation in provider documentation of enrollees' health status across provider types and groups results in natural variation and validation errors. Therefore, in the 2019 Payment Notice final rule,[23] HHS adopted the current error rate calculation methodology to evaluate material statistical deviation in failure rates. The current methodology was adopted to avoid adjusting issuers' risk scores and transfers due to expected variation and error. Instead, HHS amends an issuer's risk score only when the issuer's failure rate materially deviates from a statistically meaningful national value. HHS defines the national statistically meaningful value as the weighted mean and standard deviation of the failure rate calculated

[22] The exception to the current prospective application of HHS-RADV results is for exiting issuers, whose HHS-RADV results are applied to the risk scores and transfer amounts for the benefit year being audited. See 83 FR 16930 at 16965.

[23] See 83 FR 16930 at 16961 through 16965.

based on all issuers' HHS-RADV results. Each issuer's results are compared to these national metrics to determine whether the issuer's results are outliers. Based on outlier issuers' failure rate results, error rates are calculated and applied to outlier issuers' plan liability risk scores.[24]

Given comments received on the 2019 RADV White Paper and to help put the methodological changes proposed in this rule in context, this section outlines how the current error rate calculation methodology would apply if no changes were made since the latest policies were finalized in the 2021 Payment Notice.[25] This includes information on how HHS uses outlier issuer group failure rates to adjust enrollee risk scores, calculates an outlier issuer's error rate, and applies that error rate to the outlier issuer's plan liability risk score.

To apply the current error rate calculation methodology, HHS first uses the failure rates for each HCC to categorize all HCCs into three HCC groupings (a high, medium, or low HCC failure rate grouping). These HCC groupings are determined by first ranking all HCC failure rates and then dividing the rankings into three groupings, such that the total observations of HCCs on EDGE in each grouping are relatively equal across all issuers' initial validation audit (IVA) samples (or second validation audit (SVA) samples, if applicable), resulting in high, medium, and low HCC failure rate groupings. An issuer's HCC group failure rate is calculated as follows:

$$GFR_{G,i} = 1 - \frac{freqIVA_{G,i}}{freqEDGE_{G,i}}$$

[24] As detailed further below, these risk score changes are then used to adjust risk adjustment transfers for the applicable state market risk pool.
[25] 85 FR 29164.

Where:

$freqEDGE_{G,i}$ is the number of occurrences of HCCs in group *G* that are recorded on EDGE for all enrollees sampled from issuer *i*.

$freqIVA_{G,i}$ is the number of occurrences of HCCs in group *G* that are identified by the IVA audit (or SVA audit, as applicable) for all enrollees sampled from issuer *i*.

$GFR_{G,i}$ is issuer *i*'s group failure rate for the HCC group *G*.

HHS calculates the weighted mean failure rate and the standard deviation of each HCC group as:

$$\mu\{GFR_G\} = 1 - \frac{\sum_i freqIVA_{G,i}}{\sum_i freqEDGE_{G,i}}$$

$$Sd\{GFR_G\} = \sqrt{\frac{\sum_i \left(freqEDGE_{G,i} * \left(GFR_{G,i} - \mu\{GFR_G\}\right)^2\right)}{\sum_i freqEDGE_{G,i}}}$$

Where:

$\mu\{GFR_G\}$ is the weighted mean of $GFR_{G,i}$ of all issuers for the HCC group *G* weighted by all issuers' sample observations in each group.

$Sd\{GFR_G\}$ is the weighted standard deviation of $GFR_{G,i}$ of all issuers for the HCC group *G*.

Each issuer's HCC group failure rates are then compared to the national metrics for each HCC grouping. All enrollee HCCs identified by the IVA (or SVA, as applicable) are used to determine an issuer's failure rate for the applicable HCC group. If an issuer's failure rate for an HCC group falls outside of the 95 percent confidence interval around the weighted mean failure rate for the HCC group, that is, a failure rate further than 1.96 standard deviations from the weighted mean failure rate when assuming all issuers' group failure rates are normally

distributed, the failure rate for the issuer's HCCs in that group is considered an outlier (if the issuer meets the minimum number of HCCs for the HCC group). To calculate the outlier status thresholds, HHS calculates the lower and upper limits as:

$$LB_G = \mu\{GFR_G\} - sigma\_cutoff * Sd\{GFR_G\}$$

$$UB_G = \mu\{GFR_G\} + sigma\_cutoff * Sd\{GFR_G\}$$

Where:

$sigma\_cutoff$ is the parameter used to set the threshold for the outlier detection as the number of standard deviations away from the mean; 1.96 for a two-tailed 95 percent confidence interval as determined by a normal distribution.

$LB_G$, $UB_G$ are the lower and upper thresholds to classify issuers as outliers or not outliers for group *G*.

Outlier status is determined independently for each issuer's HCC failure rate group such that an issuer may be considered an outlier in one HCC failure rate group but not an outlier in another HCC failure rate group. Beginning with the 2019 benefit year, issuers are also not considered an outlier for an HCC group in which the issuer has fewer than 30 HCCs.[26,27] If no issuers' HCC group failure rates in a state market risk pool materially deviate from the national mean of failure rates or does not meet the minimum HCC requirements (that is, no issuers are

[26] See 85 FR 29196 – 29198.

[27] Data from issuers with fewer than 30 HCCs in an HCC group will be included in the calculation of national metrics for that HCC group, including the national mean failure rate, standard deviation, and upper and lower confidence interval bounds. Ibid.

outliers), HHS does not apply any adjustments to issuers' risk scores or to transfers in that state market risk pool.

When an issuer's HCC group failure rate is an outlier, we reduce (or increase) each of the applicable IVA sample (or SVA sample, if applicable) enrollees' HCC risk coefficients for HCCs in that group by the difference between the outlier issuer's failure rate for the HCC group and the weighted mean failure rate for the HCC group. Specifically, this will result in the sample enrollees' applicable HCC risk score components being reduced (or increased) by a partial value, or percentage, calculated as the difference between the outlier failure rate for the HCC group and the weighted mean failure rate for the applicable HCC group. Beginning with the 2019 benefit year, when the issuer meets the minimum HCC frequency requirement per an HCC group ($Freq\_EDGE_{G,i}$), this group adjustment factor $GAF_{G,i}$ amount for outliers is the distance between issuer $i$'s Group Failure Rate $GFR_{G,i}$ and the weighted mean $\mu\{GFR_G\}$. This is calculated[28] as:

If $GFR_{G,i} > UB_G$ or $GFR_{G,i} < LB_G$,

And if $Freq\_EDGE_{G,i} \geq 30$:

Then $Flag_{G,i} = "outlier"$ and $GAF_{G,i} = GFR_{G,i} - \mu\{GFR_G\}$

If $GFR_{G,i} \leq UB_G$ and $GFR_{G,i} \geq LB_G$,

Or if $Freq\_EDGE_{G,i} < 30$:

[28] This calculation sequence is printed here as it appears in the 2021 Payment Notice (85 FR 29164 at 29196-29198). In certain later sections of this proposed rule, we revised the order of similar sequences to ensure simplicity when demonstrating how the proposals in this proposed rule would be combined with the current error rate calculation methodology (including the changes finalized in the 2021 Payment Notice). The different display of these sequences does not modify or otherwise change the amendments to the outlier identification process finalized in the 2021 Payment Notice.

Then $Flag_{G,i} = "not\ outlier"$ and $GAF_{G,i} = 0$

Where:

$Flag_{G,i}$ is the indicator if the value of issuer *i*'s group failure rate for group *G* is more extreme than a calculated threshold by which we classify issuers into "outliers" or "not outliers" for group *G*.

$GAF_{G,i}$ is the calculated adjustment factor for issuer *i*'s risk score component for all sampled HCCs in group *G* that are recorded on EDGE.

The enrollee adjustment factor is then calculated by applying the group adjustment factor $GAF_{G,i}$ to individual HCCs. For example, if an issuer has one enrollee with the HIV/AIDS HCC and the issuer's HCC group adjustment rate is 10 percent (the difference between the issuer's group failure rate and the weighted mean failure rate) for the HCC group that contains the HIV/AIDS HCC, the enrollee's HIV/AIDS coefficient would be reduced by 10 percent. This reduction would be aggregated with any reductions to other HCCs for that enrollee to arrive at the overall enrollee adjustment factor. This value is calculated according to the following formula for each enrollee in stratum 1 through 9:

$$Adjustment_{i,e} = \frac{\sum_h \left(RS_{h,G,i,e} * GAF_{G,i}\right)}{\sum_h \left(RS_{h,G,i,e}\right)}$$

Where:

$RS_{h,G,i,e}$ is the risk score component of a single HCC *h* (belonging to HCC group *G*) recorded on EDGE for enrollee *e* of issuer *i*.

$Adjustment_{i,e}$ is the calculated adjustment factor to adjust enrollee *e* of issuer *i*'s EDGE risk scores.

$GAF_{G,i}$ is the calculated adjustment factor for issuer *i*'s risk score components for all sampled HCCs in group *G* that are recorded on EDGE.

The calculation of the enrollee adjustment factor above only considers risk score components related to the HCC and ignores any other risk score components (such as demographic components and RXC components). Newly identified HCCs by the IVA (or SVA as applicable) contribute to the calculation of the issuer's group failure rate but do not contribute to enrollee risk score adjustments for that enrollee and adjusted enrollee risk scores are only computed for sampled enrollees with HCCs in strata 1 through 9.

Next, for each sampled enrollee with HCCs, HHS applies the enrollee adjustment factor to each stratum 1 through 9 enrollee's risk score (including the non-HCC risk adjustment components, such as demographic components and RXC components) as recorded on the EDGE server, calculating the total adjusted enrollee risk score for these enrollees as:

$$AdjRS_{i,e} = EdgeRS_{i,e} * (1 - Adjustment_{i,e})$$

Where:

$EdgeRS_{i,e}$ is the risk score as recorded on the EDGE server of enrollee *e* of issuer *i*.

$AdjRS_{i,e}$ is the amended risk score for sampled enrollee *e* of issuer *i*.

$Adjustment_{i,e}$ is the adjustment factor by which we estimate the EDGE risk score exceeds or falls short of the initial or second validation audit projected total risk score for sampled enrollee *e* of issuer *i*.

The calculation of the total adjusted enrollee risk score $AdjRS_{i,e}$ for sample enrollees in strata 1-9 is based on the risk score recorded on EDGE server $EdgeRS_{i,e}$ that includes all risk score components (that is, both HCCs and the non-HCC components). Enrollees with no HCCs do not have enrollee adjustment factors or adjusted risk scores; however, we note that they

contribute to the calculation of the outlier issuer's group failure rate in advance of the calculation of adjustments.

After calculating the adjusted EDGE risk scores for outlier issuers' sample enrollees with HCCs, HHS calculates an outlier issuer's error rate by extrapolating the difference between the amended risk score and EDGE risk score for all enrollees (stratum 1 through 10) in the sample. The weight in the extrapolation formula associated with an enrollee's amended risk score and EDGE risk score is determined as the ratio of (1) the stratum size in the issuer's population for the enrollee's stratum, to (2) the number of sampled enrollees in the same stratum as the enrollee. Sample enrollees with no HCCs are included in the extrapolation of the error rate for outlier issuers with unchanged EDGE risk scores where $AdjRS_{i,e} = EdgeRS_{i,e}$ for enrollees with no HCCs. The formulas to compute the error rate using the stratum-weighted risk score before and after the adjustment are:

$$ErrorRate_i = 1 - \frac{\sum_e (w_{i,e} * AdjRS_{i,e})}{\sum_e (w_{i,e} * EdgeRS_{i,e})}$$

Where:

$$w_{i,e} = \frac{stratum\ size\ in\ population}{number\ of sample\ enrollees\ of\ the\ stratum}$$

Consistent with 45 CFR 153.350(c), HHS then applies the outlier issuer's error rate to adjust that issuer's applicable benefit year plan liability risk score.[29] This risk score change,

[29] Exiting outlier issuer risk score error rates are currently applied to the plan liability risk scores and risk adjustment transfer amounts for the benefit year being audited. For all other outlier issuers, risk score error rates are currently

which also impacts the state market average risk score, is then used to adjust the applicable benefit year's risk adjustment transfers for the applicable state market risk pool. Due to the budget-neutral nature of the HHS-operated program, adjustments to one issuer's risk scores and risk adjustment transfers based on HHS-RADV findings will affect other issuers in the state market risk pool (including those who were not identified as outliers) because the state market average risk score is recalculated to reflect the change in the outlier issuer's plan liability risk score. This also means that issuers that are exempt from HHS-RADV for a given benefit year may have their risk adjustment transfers adjusted based on other issuers' HHS-RADV results.

In response to stakeholder concerns, comments to the 2019 RADV White Paper, and our analyses of 2017 benefit year HHS-RADV results, HHS is proposing to modify the HCC grouping methodology used to calculate failure rates by combining certain HCCs with the same risk score coefficient for grouping purposes, and to refine the error estimation methodology to mitigate the impact of the "payment cliff" effect, in which some issuers with similar HHS-RADV findings may experience different adjustments to their risk scores and transfers. We also propose changes to mitigate the impact of HHS-RADV adjustments that result from negative error rate outlier issuers with negative failure rates.

---

applied to the plan liability risk scores and risk adjustment transfer amounts for the current transfer year. The exiting issuer exception is discussed in Section II.B.

The 2019 RADV White Paper discussed several alternatives for potential changes to HHS-RADV, and we considered those alternatives and the comments we received on them when considering which proposals to propose in this rulemaking. This proposed rule addresses only certain policies discussed in the 2019 RADV White Paper. We intend to continue to analyze HHS-RADV results and consider potential further refinements to the HHS-RADV methodology for future benefit years.

1. HCC Grouping for Failure Rate Calculation

HHS groups medical conditions in multiple distinct ways during the risk adjustment and HHS-RADV processes. These grouping processes include:

For risk adjustment model development:

(1) The hierarchies of Hierarchical Condition Categories (HCCs),

(2) HCC coefficient estimation groups,[30]

(3) A priori stability constraints, and

(4) Hierarchy violation constraints.

And, for HHS-RADV:

(5) HHS-RADV HCC failure rate groups.

The first four of these grouping processes are related to the development and estimation of coefficients in the HHS risk adjustment models, while the fifth is related to error estimation during HHS-RADV. These grouping processes are not concurrent. The grouping processes

[30] The current HCC coefficient estimation groups for the adult models are identified in Column B of Table 6 in the "Do It Yourself" Software. The current HCC coefficient estimation groups for the child models are identified in Column B of Table 7 in the "Do it Yourself" Software.

related to the risk adjustment models are implemented prior to the benefit year and interact with HHS-RADV HCC failure rate groups that are implemented after the benefit year. Our experience in the initial years of HHS-RADV found that differences among the risk adjustment and HHS-RADV grouping procedures interact in varying ways and may result in greater or lesser HHS-RADV adjustments than may be warranted in certain circumstances. Examples of these interactions are discussed later in this proposed rule.

The first grouping of medical conditions –HCCs – is used to aggregate thousands of standard disease codes into medically meaningful but statistically manageable categories. HCCs in the 2019 benefit year HHS risk adjustment models were derived from ICD-9-CM codes[31] that are aggregated into diagnostic groups (DXGs), which are in turn aggregated into broader condition categories (CCs). Then, clinical hierarchies are applied to the CCs, so that an enrollee receives an increase to their risk score for only the most severe manifestation among related diseases that may appear in their medical claims data on an issuer's EDGE server.[32] Condition categories become Hierarchical Condition Categories (HCCs) once these hierarchies are imposed.

[31] In the 2021 Payment Notice, we finalized several updates to the HHS-HCC clinical classification by using more recent claims data to develop updated risk factors that apply beginning with the 2021 benefit year risk adjustment models. See 85 FR 29164 at 29175 (May14, 2020). Also see The Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program (June 17, 2019) (2019 HHS-HCC Potential Updates Paper), available at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf.

[32] The process for creating hierarchies is an iterative process that considers severity, as well as costs of the HCCs in the hierarchies and clinical input, among other factors. For information on this process, see section 2.3 of the 2019 HHS-HCC Potential Updates Paper.

As noted above, for a given hierarchy, if an enrollee has more than one HCC recorded in an issuer's EDGE server, only the most severe of those HCCs will be applied for the purposes of risk adjustment model and plan liability risk score calculation.[33] For example, respiratory distress diagnosis codes are organized in a hierarchy consisting of three HCCs arranged in descending order of clinical severity from (1) HCC 125 Respirator Dependence/Tracheostomy Status to (2) HCC 126 Respiratory Arrest to (3) HCC 127 Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes. An enrollee may have diagnosis codes in two respiratory distress HCCs, but once hierarchies are imposed, that enrollee would only be assigned the single highest severity HCC in the hierarchy. Thus, an enrollee with diagnosis codes in HCC 126 Respiratory Arrest and HCC 127 Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes would only be assigned the single highest HCC (in this case, HCC 126 Respiratory Arrest). Although HCCs reflect hierarchies among related disease categories, for unrelated diseases, multiple HCCs can accumulate for those enrollees, that is, the model is "additive." For example, an enrollee with both diabetes and asthma would have (at least) two separate HCCs coded and the predicted cost for that enrollee will reflect increments for both conditions.

In the risk adjustment models, estimated coefficients of the various HCCs within a hierarchy will ensure that more severe and expensive HCCs within that hierarchy receive higher risk factors than less severe and less expensive HCCs. Additionally, as a part of the recalibration of the risk adjustment models, HHS has grouped some HCCs so that the coefficients of two or

[33] Once hierarchies are imposed, CC code groups are referred to as HCCs.

more HCCs are equal in the fitted risk adjustment models and only one model factor is assigned to an enrollee regardless of the number of HCCs from that group present for that enrollee on the issuer's EDGE server,[34] giving rise to the second set of condition groupings used in risk adjustment. We impose these HCC coefficient estimation groups for a number of reasons, including the limitation of diagnostic upcoding by severity within an HCC hierarchy and the reduction of additivity within disease groups (but not across disease groups) in order to decrease the sensitivity of the models to coding proliferation.

Some of these HCC coefficient estimation groups occur within hierarchies. For example, HCC 126 Respiratory Arrest and HCC 127 Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes within the respiratory distress hierarchy are grouped into a single HCC coefficient estimation group. However, some HCC coefficient estimation groups include HCCs that do not share a hierarchy. For example, another HCC coefficient estimation group consists of HCC 61 Osteogenesis Imperfecta and Other Osteodystrophies and HCC 62 Congenital/Developmental Skeletal and Connective Tissue Disorders. Within an HCC coefficient estimation group, each HCC will have the same coefficient in our risk adjustment models. However, as with hierarchies, only one risk marker is triggered by the presence of one or more HCCs in the HCC coefficient estimation groups. These HCC coefficient estimation groups are identified in DIY Software Table 6 for the adult models and DIY Software Table 7 for the

[34] As described in the June 17, 2019 document "Potential Updates to HHS-HCCs for the HHS-operated Risk Adjustment Program", available at https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf#page=11.

child models. The adult model HCC coefficient estimation groups for the V05 risk adjustment models[35] are displayed in Table 1:

**Table 1: HCC Coefficient Estimation Groups from Adult Risk Adjustment Models V05**

| HHS HCC | V05 HHS-HCC LABEL | Adult Model HCC Coefficient Estimation Group |
|---|---|---|
| 19 | Diabetes with Acute Complications | G01 |
| 20 | Diabetes with Chronic Complications | G01 |
| 21 | Diabetes without Complication | G01 |
| 26 | Mucopolysaccharidosis | G02A |
| 27 | Lipidoses and Glycogenosis | G02A |
| 29 | Amyloidosis, Porphyria, and Other Metabolic Disorders | G02A |
| 30 | Adrenal, Pituitary, and Other Significant Endocrine Disorders | G02A |
| 54 | Necrotizing Fasciitis | G03 |
| 55 | Bone/Joint/Muscle Infections/Necrosis | G03 |
| 61 | Osteogenesis Imperfecta and Other Osteodystrophies | G04 |
| 62 | Congenital/Developmental Skeletal and Connective Tissue Disorders | G04 |
| 67 | Myelodysplastic Syndromes and Myelofibrosis | G06 |
| 68 | Aplastic Anemia | G06 |
| 69 | Acquired Hemolytic Anemia, Including Hemolytic Disease of Newborn | G07 |
| 70 | Sickle Cell Anemia (Hb-SS) | G07 |
| 71 | Thalassemia Major | G07 |
| 73 | Combined and Other Severe Immunodeficiencies | G08 |
| 74 | Disorders of the Immune Mechanism | G08 |
| 81 | Drug Psychosis | G09 |
| 82 | Drug Dependence | G09 |
| 106 | Traumatic Complete Lesion Cervical Spinal Cord | G10 |

[35] The shorthand "V05" refers to the current HHS-HCC classification for the HHS risk adjustment models, which applies through the 2020 benefit year.

| HHS HCC | V05 HHS-HCC LABEL | Adult Model HCC Coefficient Estimation Group |
|---|---|---|
| 107 | Quadriplegia | G10 |
| 108 | Traumatic Complete Lesion Dorsal Spinal Cord | G11 |
| 109 | Paraplegia | G11 |
| 117 | Muscular Dystrophy | G12 |
| 119 | Parkinson's, Huntington's, and Spinocerebellar Disease, and Other Neurodegenerative Disorders | G12 |
| 126 | Respiratory Arrest | G13 |
| 127 | Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes | G13 |
| 128 | Heart Assistive Device/Artificial Heart | G14 |
| 129 | Heart Transplant | G14 |
| 160 | Chronic Obstructive Pulmonary Disease, Including Bronchiectasis | G15 |
| 161 | Asthma | G15 |
| 187 | Chronic Kidney Disease, Stage 5 | G16 |
| 188 | Chronic Kidney Disease, Severe (Stage 4) | G16 |
| 203 | Ectopic and Molar Pregnancy, Except with Renal Failure, Shock, or Embolism | G17 |
| 204 | Miscarriage with Complications | G17 |
| 205 | Miscarriage with No or Minor Complications | G17 |
| 207 | Completed Pregnancy With Major Complications | G18 |
| 208 | Completed Pregnancy With Complications | G18 |
| 209 | Completed Pregnancy with No or Minor Complications | G18 |

The HHS-HCC model also incorporates a small number of "a priori stability constraints" to stabilize estimates that might vary greatly due to small sample size.[36] These a priori stability

[36] For example, we previously finalized a constraint for six coefficients associated with seven transplant status HCCs (excluding kidney transplants) in the child model, as the sample sizes of transplants are smaller in the child than the adult model. Because the levels and changes in the child transplant relative coefficients appeared to be dominated by random instability at the time, we believed the accuracy of the models were improved by constraining these coefficients. See the HHS Notice of Benefit and Payment Parameters for 2016, Final Rule, 80 FR 10749 at 10761 (February 27, 2015).

constraints differ from the HCC coefficient estimation groups in how the corresponding estimates are counted. In contrast to HCC coefficient estimation groups, with a priori stability constraints, a person can have more than one indicated condition (each with the same coefficient value) as long as the HCCs are not in the same hierarchy. As seen in Table 2, prior to the 2021 benefit year recalibration,[37] only one a priori stability constraint was applied to the models, and this constraint was only applied to the child models.

**Table 2: HCCs Subject to A Priori Stability Constraints in Risk Adjustment Child Models V05**

| HHS HCC | V05 HHS-HCC LABEL | Child Model *A Priori* Stability Constraint |
|---|---|---|
| 18 | Pancreas Transplant Status/Complications | S1 |
| 34 | Liver Transplant Status/Complications | S1 |
| 41 | Intestine Transplant Status/Complications | S1 |
| 128 | Heart Assistive Device/Artificial Heart | S1 |
| 129 | Heart Transplant | S1 |
| 158 | Lung Transplant Status/Complications | S1 |
| 251 | Stem Cell, Including Bone Marrow, Transplant Status/Complications | S1 |

HCC coefficient estimation group constraints and a priori stability constraints are both applied in the initial phase of risk adjustment regression modeling. Other constraints may be applied in later stages depending on regression results. For example, HCCs may be constrained equal to each other if there is a hierarchy violation (a lower severity HCC has a higher estimate

[37] In the 2021 Payment Notice (85 FR 29164 at 29178), we introduced an additional a priori stability constraint to the child risk adjustment models, constraining HCC 218 Extensive Third Degree Burns and HCC 223 Severe Head Injury to have the same risk adjustment coefficient due to small sample size. We also revised the current single transplant stability constraint in the child models (shown in Table 2) into two stability constraints to better distinguish transplant cost differences.

than a higher severity HCC in the same hierarchy).[38] HCC coefficients may also be constrained to 0 if the estimates fitted by the regression model are negative.

The final set of groupings is imposed during the error estimation stage of the HHS-RADV process. In this process, HCCs are categorized into low, medium, and high HCC failure rate groups. These groupings are designed to balance the need to assess the impact of medical coding errors of individual HCCs on risk scores and risk adjustment transfers and the need to assess failure rates on enough HCCs to provide statistically meaningful HHS-RADV results. Furthermore, these groupings are intended to reflect the fact that some HCCs are more difficult to code accurately than other HCCs and to provide national standards that take into account the level of coding difficulty for a given HCC.

To create the HHS-RADV HCC failure rate groupings, the first step is to calculate the national average failure rate for each HCC individually. The second step involves ranking HCCs in order of their failure rates and then dividing them into three groups — a low, medium, and high failure rate group — such that the total counts of HCCs in each group nationally as recorded in EDGE data across all IVA samples (or SVA samples if applicable) are roughly equal. These HCC failure rate groups form the basis of the failure rate outlier determination process, with each failure rate group receiving an independent assessment of outlier status for each issuer.[39]

---

[38] For example, in the 2019 benefit year of risk adjustment adult models, HCC 88 (Major Depression and Bipolar Disorders) and HCC 89 (Reactive and Unspecified Psychosis, Delusional Disorders) were constrained to be equal due to a hierarchy violation occurring. Therefore, these HCCs in the 2019 benefit year final adult models have the same risk scores; however, these two HCCs are not grouped (as shown in Table 6, Column B of 2019 benefit year DIY Software).

[39] For a table of the HCC failure rate groupings for 2017 benefit year HHS-RADV, see the 2019 RADV White Paper, Appendix E.

Based on our experience with the initial years of HHS-RADV, HHS observed that, in certain situations, the risk adjustment HCC hierarchies and HCC coefficient estimation groups can influence and interact with the HHS-RADV HCC failure rate groupings in varying ways that could result in misalignments.[40] For example:

- Scenario 1: HCCs in the same HCC hierarchy with different coefficients are sorted into different HHS-RADV HCC failure rate groupings.

++ If one HCC is commonly miscoded as another HCC in the same hierarchy, but the two HCCs are sorted into different HCC failure rate groupings in HHS-RADV, an issuer may be flagged as an outlier in either of the HCC failure rate groupings where one HCC is missing or the other HCC is newly found.

++ For example, HCC 8 Metastatic Cancer and HCC 11 Colorectal, Breast (Age < 50), Kidney, and Other Cancers are in the same hierarchy in risk adjustment, but for the 2017 benefit year of HHS-RADV, HCC 8 was in the medium HCC failure rate grouping and HCC 11 was in the high HCC failure rate grouping. In validating an enrollee with HCC 8 in HHS-RADV, the IVA or SVA Entity may find that an enrollee with HCC 8 reported in EDGE is not validated as having HCC 8, which is at the top of the HCC hierarchy in risk adjustment, but the enrollee may have been found to have HCC 11 in the issuer's HHS-RADV audit data. In this case, HCC 8 would be considered missing in the medium HCC failure rate grouping, and HCC 11 would be considered found in the high HCC failure rate grouping.

[40] See Section 3.3 of the 2019 RADV White Paper.

++ This circumstance would influence the failure rate for that issuer, potentially leading to the issuer being classified as an outlier in an HCC failure rate grouping. If the issuer is found to be an outlier in one of the two failure rate groupings, the issuer's HCC failure rate would not represent the actual difference in risk and costs between these two coefficients.

- Scenario 2: HCCs in the same HCC hierarchy with different coefficients are sorted into the same HHS-RADV HCC failure rate grouping.

++ If one HCC is commonly miscoded as another HCC in the same hierarchy, and the two HCCs are sorted into the same HCC failure rate grouping, the issuer may not be flagged as an outlier for that HCC grouping. This may occur because the failure to validate an HCC and the discovery of a new HCC in that same HCC failure rate grouping have a net impact of zero on the total final value of the issuer's failure rate. For purposes of the calculation of the failure rate, there would appear to be no difference between the two HCCs, even though they have different coefficients in risk adjustment.

++ For example, HCC 35 End-Stage Liver Disease and HCC 34 Liver Transplant Status/Complications are in the same hierarchy in risk adjustment and were both sorted into the medium HCC failure rate grouping in the 2017 benefit year HHS-RADV results. In validating an enrollee with HCC 35 in HHS-RADV, the IVA or SVA Entity may find that an enrollee with HCC 35 reported in EDGE is not validated as having HCC 35, but the enrollee may have been found to have HCC 34 in issuer's HHS-RADV audit data. In this case, not validating HCC 35 and finding HCC 34 in the same HCC grouping in HHS-RADV would, when taken together, have no net impact on the issuer's HCC group failure rate.

++ This situation would influence the failure rate for that issuer, potentially leading to the issuer not being classified as an outlier in an HCC failure rate grouping even though the two

HCCs have different risk and costs. If the issuer is not found to be an outlier in the applicable failure rate grouping, the issuer's HHS-RADV adjustment would not represent the actual difference in risk and costs between these two coefficients.

- Scenario 3: HCCs in the same HCC coefficient estimation group are sorted into different HCC failure rate groupings.

++ In this situation, a miscoding of one HCC for the other may lead to the issuer being identified as a positive outlier in one HCC failure rate grouping or a negative outlier in another, despite there being no difference in risk score due to the coding error.

++ For example, HCC 54 Necrotizing Fasciitis and HCC 55 Bone/Joint/Muscle Infections/Necrosis share a hierarchy and an HCC coefficient estimation group in risk adjustment, resulting in risk score coefficients constrained to be equal, but for 2017 benefit year HHS-RADV, HCC 54 was in the high failure rate HCC grouping, while HCC 55 was in the medium failure rate HCC grouping. In validating an enrollee with HCC 54 in HHS-RADV, the IVA or SVA Entity may find that an enrollee with HCC 54 reported in EDGE is not validated as having HCC 54, but the enrollee may have been found to have HCC 55 in issuer's HHS-RADV audit data.

++ In this case, when taken together with the issuer's other HHS-RADV results, HCCs in the same HCC coefficient estimation group could contribute to an issuer's failure rate in a HCC failure rate grouping, even though the HCCs do not have different risk scores and an adjustment to risk scores is not conceptually warranted. If the issuer is found to be an outlier in one of the two failure rate groupings, the issuer's HCC failure rate would not represent actual differences in risk or costs between these two coefficients.

Based on HHS's initial analysis of the occurrence of these scenarios in the 2017 benefit year HHS-RADV results,[41] and in response to comments to the 2019 RADV White Paper, HHS is considering an option in this proposed rule to address the influence of the HCC hierarchies and HCC coefficient estimation groups on the HCC failure rate groupings in HHS-RADV. Our intention is to address this issue on an interim basis while we continue to assess different longer-term options, including potential significant changes to the outlier determination process, which require additional analysis and consideration before proposing.

To address Scenario 3, we propose to modify the creation of HHS-RADV HCC failure rate groupings and place all HCCs that share an HCC coefficient estimation group in the adult risk adjustment models (see Table 1 for the list of the HCC coefficient estimation groups in the V05 classification) into the same HCC failure rate grouping. Specifically, we propose that when HHS calculates EDGE and IVA frequencies for each individual HCC and prior to sorting the HCCs into low, medium, and high failure rate groups for HHS-RADV, HCCs that are in the same HCC coefficient estimation group in the adult risk adjustment models (and, therefore, have coefficients constrained to be equal to one another) would be aggregated into one HCC. These

[41] As discussed in the 2019 RADV White Paper, we performed an initial review of the occurrence of these scenarios in the 2017 benefit year HHS- RADV results. Of all the HCCs in EDGE that were not validated in the audit data, about 1/8th represented HCCs that IVA or SVA auditors coded as different HCCs within the same hierarchy. Of the HCCs that were newly found in the audit data – that is, they were not recorded in the original EDGE data – around 1/3rd represented HCCs that were newly found because they were originally reported on EDGE as a different HCC in the same hierarchy. However, we note that these occurrences reflect both HCCs sorted into different HCC failure rate groups and HCCs sorted into the same HCC failure rate groups, including a scenario, discussed in the whitepaper wherein HCCs in the same hierarchy and the same HCC coefficient estimation group are sorted into the same HCC failure rate group, which would have no impact on failure rate and would not warrant any adjustment to risk score. Therefore, for many issuers, these occurrences would be unlikely to impact whether they were an outlier in an HCC failure rate grouping. However, we note that the initial review discussed in the white paper did not consider HCCs that share an HCC coefficient estimation group, but do not share a hierarchy.

new frequencies, including the aggregated frequencies of HCC coefficient estimation groups and the frequencies of all other unconstrained HCCs, treated separately, would be considered frequencies of "Super HCCs".

In the current process,[42] before sorting into the three HCC failure rate groups, failure rates for each HCC are calculated individually as:

$$FR_h = 1 - \frac{freqIVA_h}{freqEDGE_h}$$

Where:

*h* is the index of the $h^{th}$ HCC code;

$freqEDGE_h$ is the frequency of an HCC *h* occurring in EDGE data; that is, the number of sampled enrollees recording HCC *h* in EDGE data across all issuers participating in HHS-RADV;

$freqIVA_h$ is the frequency of an HCC *h* occurring in IVA results (or SVA results, as applicable); that is, the number of sampled enrollees recording HCC *h* in IVA (or SVA, as applicable) results across all issuers participating in HHS-RADV; and

$FR_h$ is the national overall (average) failure rate of HCC *h* across all issuers participating in HHS-RADV.

In the proposed methodology, this step would be modified as:

$$FR_c = 1 - \frac{freqIVA_c}{freqEDGE_c}$$

[42] See the 2018 HHS-RADV protocols, section 11.3.1, available at: https://www.regtap.info/uploads/library/HRADV_2018Protocols_070319_5CR_070519.pdf

Where:

$c$ is the index of the $c^{th}$ Super HCC;

$freqEDGE_c$ is the frequency of a Super HCC $c$ occurring in EDGE data across all issuers participating in HHS-RADV; that is, the sum of $freqEDGE_h$ for all HCCs that share an HCC coefficient estimation group in the adult models:

$$freqEDGE_c = \sum_h freqEDGE_{h,c}$$

When an HCC is not in an HCC coefficient estimation group in the adult risk adjustment models, the $freqEDGE_c$ for that HCC will be equivalent to $freqEDGE_h$;

$freqIVA_c$ is the frequency of a Super HCC $c$ occurring in IVA results (or SVA results, as applicable) across all issuers participating in HHS-RADV; that is, the sum of $freqIVA_h$ for all HCCs that share an HCC coefficient estimation group in the adult risk adjustment models:

$$freqIVA_c = \sum_h freqIVA_{h,c}$$

And;

$FR_c$ is the national overall (average) failure rate of Super HCC $c$ across all issuers participating in HHS-RADV.

Then, the failure rates for all Super HCCs, both those composed of a single HCC and those composed of the aggregate frequencies of HCCs that share an HCC coefficient estimation group in the adult risk adjustment models, would be grouped according to the current HHS-RADV failure rate grouping methodology.

As an illustrative example, this proposal would mean that, for purposes of HHS-RADV groupings, two of the three current respiratory distress HCCs in the adult risk adjustment models, HCC 126 Respiratory Arrest and HCC 127 Cardio-Respiratory Failure and Shock, Including Respiratory Distress Syndromes, would be aggregated into one Super HCC because they have the same estimated costs and share an HCC coefficient estimation group. That Super HCC would then be sorted into a failure rate group according to its overall national failure rate. As such, all validations or failures to validate either of the two HCCs composing the Super HCC would contribute to the failure rate for the same HCC failure rate grouping. However, if an enrollee with one of the two HCCs in the Super HCC reported on EDGE was not validated as having the EDGE reported HCC but is found to have the other HCC in the Super HCC (e.g., an enrollee with HCC 126 reported on EDGE is not validated as having HCC 126 but is found to have HCC 127), the issuer's failure rate would not be affected. This approach would ensure that HCCs with the same estimated costs in the adult risk adjustment models that share an HCC coefficient estimation group do not contribute to an issuer's failure rate in a HCC failure rate grouping. To promote fairness and ensure the integrity of the program, we do not believe that issuers should be considered to have an HHS-RADV error for similar conditions from the same HCC coefficient estimation group and, as a result, were estimated as having the same risk in the adult risk adjustment models. This proposal to aggregate the frequencies of HCCs in the same HCC coefficient estimation group in the adult risk adjustment models would refine the HHS-RADV methodology to better identify and focus outlier determinations on actual differences in risk and costs. Based on our testing of this proposed policy on 2017 benefit year HHS-RADV results, we estimate that by creating the proposed Super HCCs, approximately 98.1 percent of the occurrences of HCCs on EDGE belong to HCCs that would be assigned to the same failure

rate groups under the proposed methodology as they have been under the current methodology as seen in Table 3. Although the impact on individual issuer results may vary depending upon the accuracy of their initial data submissions and the rate of occurrence of various HCCs in their enrollee population, the national metrics used for HHS-RADV would only be slightly affected, as seen in Table 4. The stability of these metrics and high proportion of EDGE frequencies of HCCs that would be assigned to the same failure rate group under the proposed and current sorting methodologies reflects that the most common conditions will have similar failure rates if this proposal is adopted. However, the failure rate estimates of less common conditions may be stabilized with the proposed creation of Super HCCs by ensuring these conditions are grouped alongside more common, related conditions.

In testing this proposal to create the Super HCCs in HHS-RADV, we grouped HCCs in the same HCC coefficient estimation group in the adult risk adjustment models. To do this, we used variables in Column B in Table 6 of the HHS-Developed Risk Adjustment Model Algorithm "Do It Yourself" software[43] to determine the candidate HCCs that should be incorporated into Super HCCs under this policy proposal. If a set of candidate HCCs are all from the same HCC coefficient estimation group, they would be grouped into one Super HCC in HHS-RADV. Each remaining HCC that does not meet these criteria would be assigned to its own Super HCC prior to determining the HCC failure rate grouping. We chose to use the adult risk adjustment models for testing because the majority of the population with HCCs in the

[43] 2017 Benefit Year Risk Adjustment: HHS-Developed Risk Adjustment Model Algorithm "Do It Yourself (DIY)" Software. Technical Details. July 21, 2017. Assessed at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/DIY-Tables-7-12-2017.xlsx.

HHS-RADV samples are subject to the adult models (88.3 percent for the 2017 benefit year).[44] As such, the adult models' HCC coefficient estimation groups will be applicable to the vast majority of enrollees and we believe that the use of HCC coefficient estimation groups present in the adult risk adjustment models sufficiently balances the representativeness and precision of HCC failure rate estimates across the entire population in aggregate and may be used as the source for the proposed creation of Super HCCs for all RADV sample enrollees, regardless of the risk adjustment model to which they are subject.

In developing this policy, we limited the grouping of risk adjustment HCCs into Super HCCs for HHS-RADV to HCC coefficient estimation groups alone and have not considered including a priori stability constraints or hierarchy violation constraints in the aggregation of Super HCCs. A priori stability constraints currently are only applied to a limited number of HCCs in the child models and are applied differently than HCC hierarchies and HCC coefficient estimation groups. Whereas enrollees can only receive one HCC from a hierarchy or one model factor from a coefficient estimation group (for example, one factor for the presence of either HCC 61 Osteogenesis Imperfecta and Other Osteodystrophies or HCC 62 Congenital/Developmental Skeletal and Connective Tissue Disorders), enrollees may receive more than one HCC when there is an a priori stability constraint (for example, HCC 129 Heart Transplant and HCC 158 Lung Transplant Status/Complications in the child model). Although HCCs subject to a priori stability constraints will have the same coefficient value, the possible

[44] This was calculated after removing issuers in Massachusetts and incorporating cases where issuers failed pairwise and the SVA sub-sample was used.

additive nature of these HCCs suggests that a failure to validate one HCC subject to an a priori stability constraint paired with the IVA or SVA entity identifying a different HCC subject to the same a priori stability constraint does not constitute a swapping of HCCs in the same way that a similar scenario among HCCs in a common HCC coefficient estimation group would. As such, we do not find it necessary or appropriate to include a priori stability constraints in the aggregation of Super HCCs.

We also did not consider hierarchy violation constraints as a part of the sorting algorithm in order to balance complexity and consistency, as hierarchy violation constraints in the risk adjustment models can change from year-to-year as a natural result of risk adjustment model coefficient annual recalibration updates. These year-to-year changes would make HCC groupings for these HCCs less stable and transparent, and would reduce predictability for issuers.

For the above mentioned reasons, we propose to combine HCCs in HCC coefficient estimation groups in the adult risk adjustment models into Super HCCs prior to sorting the HCCs into low, medium and high failure rate groups for HHS-RADV, starting with the 2019 benefit year of HHS-RADV. If finalized as proposed, these Super HCC groupings would apply to all RADV sample enrollees, regardless of the risk adjustment models to which they are subject. Once sorted into failure rate groups, the failure rates for all Super HCCs, both those composed of a single HCC and those composed of the aggregate frequencies of HCCs that share an HCC coefficient estimation group in the adult risk adjustment models, would be grouped according to the current HHS-RADV failure rate grouping methodology.

We solicit comment on all aspects of this proposal. In particular, we solicit comments on the proposed use of the HCC coefficient estimation groups to identify the HCCs that would be aggregated into Super HCCs in HHS-RADV and whether we should also consider incorporating

a priori stability constraints from the child models, or hierarchy violation constraints from the adult risk adjustment models as part of HHS-RADV Super HCCs. We also solicit comment on whether, in addition to the Super HCCs based on the adult risk adjustment models, CMS should create separate infant Super HCCs for each severity type in the infant risk adjustment models. As we considered with the adult risk adjustment model-based Super HCCs, if we were to adopt separate infant model-based Super HCCs, we solicit comments on whether we should incorporate only the HCC coefficient groupings inherent in the infant severity level determination process, or both these groupings and any hierarchy violation constraints that may occur in the infant models. The latter option may make the composition of HCC groups less stable year-to-year, but may more comprehensively address Scenario 3 when it occurs and reflect the full risk structure of HCC hierarchies as expressed in infant risk adjustment models.

Additionally, we solicit comment regarding the impact of COVID-19 on the proposed changes to the HCC grouping methodology for error rate calculation. In particular, we solicit comment on whether the need for providers to focus on caring for patients during the COVID-19 pandemic could impact the completeness of the data that would be used to implement the new HCC grouping methodology for HHS-RADV, such that we should consider a later applicability date if we finalize this proposal.

**Table 3: Estimated Proposed Changes in the HCC Groupings Using Super HCCs Based on Adult Model HCC Coefficient Estimation Groups (Using the 2017 Benefit Year HHS-RADV Results)**

| Count of HCC Categories in Each Failure Rate Group | | Super HCCs using HCC Coefficient Estimation Groups (Proposed Option) | | |
|---|---|---|---|---|
| | | Low | Medium | High |
| Current Methodology | Low | 31 | 1 | 1 |
| | Medium | 2 | 29 | 4 |

| | | | | |
|---|---|---|---|---|
| | High | 1 | 5 | 53 |
| **Frequency of HCC Occurrence on EDGE** | | **Super HCCs using HCC Coefficient Estimation Groups (Proposed Option)** | | |
| | | Low | Medium | High |
| Current Methodology | Low | 32.2% | 0.0% | 0.0% |
| | Medium | 0.1% | 33.0% | 1.0% |
| | High | 0.3% | 0.3% | 32.9% |

**Table 4: Estimated Proposed National Metrics in the HCC Groupings Using Super HCCs Based on Adult Model HCC Coefficient Estimation Groups (Using the 2017 Benefit Year HHS-RADV Results)**

| HCC Grouping Options | Group | Weighted Mean Failure Rate | Weighted Std. Dev | Lower Threshold | Upper Threshold |
|---|---|---|---|---|---|
| Current | Low | 0.0476 | 0.0973 | -0.1431 | 0.2382 |
| | Med | 0.1549 | 0.0992 | -0.0395 | 0.3493 |
| | High | 0.2621 | 0.1064 | 0.0536 | 0.4706 |
| Super HCCs using HCC Coefficient Estimation Groups (Proposed Option) | Low | 0.0496 | 0.0959 | -0.1384 | 0.2376 |
| | Med | 0.1557 | 0.0994 | -0.0392 | 0.3506 |
| | High | 0.2595 | 0.1065 | 0.0508 | 0.4682 |

2. "Payment cliff" effect

The HHS-RADV error rate calculation methodology is based on the identification of outliers, as determined using certain national thresholds. In the case of the current error rate calculation methodology, those thresholds are used to determine whether an issuer is an outlier, and to determine the error rate that will be used to adjust risk scores. As previously discussed, under the current methodology, 1.96 standard deviations on both sides of the confidence interval around the weighted HCC group means are the thresholds currently used to determine whether an issuer is an outlier. In practice, these thresholds mean that an issuer with failure rates outside the 1.96 standard deviations range for any of the HCC failure groups is deemed an outlier and

receives an adjustment to its risk score, while an issuer with failure rates inside the 1.96 standard deviations range for all groups receives no adjustment to its risk score.[45]

As stated in the 2021 Payment Notice, beginning with the 2019 benefit year, when the issuers meets the minimum HCC requirement per an HCC group ($Freq\_EDGE_{G,i}$), the group adjustment factor for outliers is the distance between issuer *i*'s Group Failure Rate $GFR_{G,i}$ and the weighted mean $\mu\{GFR_G\}$, calculated[46] as:

If $GFR_{G,i} > UB_G$ or $GFR_{G,i} < LB_G$:

And if $Freq\_EDGE_{G,i} \geq 30$:

Then $Flag_{G,i} = "outlier"$ and $GAF_{G,i} = GFR_{G,i} - \mu\{GFR_G\}$

If $GFR_{G,i} \leq UB_G$ and $GFR_{G,i} \geq LB_G$,

Or if $Freq\_EDGE_{G,i} < 30$:

Then $Flag_{G,i} = "not\ outlier"$ and $GAF_{G,i} = 0$

Where:

$Flag_{G,i}$ is the indicator if issuer *i*'s group failure rate for group *G* is located beyond a calculated threshold that we use to classify issuers into "outliers" or "not outliers" for group *G*.

[45] An issuer with no error rate would not have its risk score adjusted due to HHS-RADV, but that issuer may have its risk adjustment transfer impacted if there is another issuer(s) in the state market risk pool that is an outlier.

[46] This calculation sequence is printed here as it appears in the 2021 Payment Notice (85 FR 29164 at 29196-29198). In later sections of this rule, we revised the order of similar sequences for simplicity when demonstrating how this sequence would be combined with proposals in this proposed rule. The different display does not modify or otherwise change the amendments to the outlier identification process finalized in the 2021 Payment Notice.

$GAF_{G,i}$ is the calculated adjustment factor to adjust issuer *i*'s EDGE risk score components for all sampled HCCs in group *G*.

For each sampled enrollee with HCCs, the group adjustment factor (GAF) is applied at the individual HCC level to all EDGE HCCs in the HCC grouping in which the issuer is an outlier. For example, if an issuer's sample has one enrollee with the HIV/AIDS HCC and the issuer's HCC GAF[47] is 10 percent (the difference between the outlier issuer's group failure rate and the weighted mean group failure rate) for the HCC group that contains the HIV/AIDS HCC, the enrollee's HIV/AIDS HCC risk score coefficient would be reduced by 10 percent. This reduction would be aggregated with any reductions to other HCCs for that enrollee to arrive at the overall enrollee adjustment factor for each sample enrollee in stratum 1 through 9. Next, each stratum 1 through 9 sample enrollee's enrollee adjustment factor is applied to that enrollee's entire EDGE risk score (including the non-HCC risk adjustment components) to calculate an adjusted risk score for that sample enrollee. These adjusted risk scores are extrapolated to the issuer's population strata and aggregated with the unadjusted risk scores of stratum 10 enrollees in the calculation of the issuer's error rate.

Some stakeholders have expressed concern that the failure rates of issuers that are just outside of the confidence intervals receive an adjustment, even though they may not be significantly different from the failure rates of issuers just inside the confidence intervals who receive no adjustment, creating a "payment cliff" or "leap frog" effect. For example, an issuer

[47] To more clearly distinguish between the enrollee adjustment factor and the group adjustment factor, for the purposes of this proposed rule, we use GAF instead of "adjustment".

with a low HCC group failure rate of 23.9 percent would be considered a positive error rate outlier for that HCC group based on the 2017 benefit year national failure rate statistics, because the upper bound confidence interval for the low HCC group is 23.8 percent. That issuer's GAF would be calculated based on the difference between the weighted low HCC group mean of 4.8 percent and the issuer's 23.9 percent failure rate for that HCC group. Under this example, the issuer's GAF would be 19.1 percent, and that GAF would be applied to the enrollee-level risk score coefficients for enrollees in the issuer's sample who have HCCs in the HCC failure rate group for which the issuer was determined to be an outlier. At the same time, another issuer with a low HCC group failure rate of 23.7 percent would receive no adjustment to its risk score as a result of HHS-RADV. While this result is due to the nature of establishing and using a threshold, some stakeholders have recommended mitigating this effect by calculating error rates based on the position of the bounds of the confidence interval for the HCC group and not on the position of the weighted mean for the HCC group. Others have recommended not adjusting issuers' risk scores in the case of negative error rate issuers to limit the impact of these adjustments on issuers who are not determined to be outliers.[48]

As we have previously discussed,[49] we have concerns about only adjusting issuers' risk scores for positive error rate outliers. However, we recognize that changing the calculation and application of an outlier issuer's error rate may be appropriate if the outlier issuer is not statistically different from the issuers within the confidence intervals. Therefore, to promote

[48] See Section II.A.3 for proposals intended to mitigate the impact of HHS-RADV adjustments for negative error rate issuers with negative failure rates.
[49] See, for example, Section 4.4.3 of the 2019 RADV White Paper. Also see 84 FR 17504 through 17508.

fairness, HHS's focus in considering potential changes to mitigate the payment cliff in the calculation of error rates is on situations where issuers with failure rates that are close to the bounds of the confidence intervals are not substantially different from issuers with failure rates inside the confidence intervals. To address this issue, we are considering potential modifications to the error rate calculation that maintain the two-sided approach of HHS-RADV through which both positive and negative error rate outliers would continue to receive risk score adjustments.

While HHS considered several possible methods to address the payment cliff in the 2019 RADV White Paper, we are proposing to address the payment cliff by adding a sliding scale adjustment to the current error rate calculation, such that different adjustments would be applied to issuers based on their distance from the mean and the farthest outlier threshold. This proposed approach would employ additional thresholds to create a smoothing of the error rate calculation beyond what the current methodology allows and to help reduce the disparity of risk score adjustments using a linear adjustment.[50] We are proposing to make this modification beginning with 2019 benefit year HHS-RADV.

To apply the sliding scale adjustment, we propose to modify the calculation of the GAF by providing a linear sliding scale adjustment, for issuers whose failure rates are near the point at which the payment cliff occurs. For those issuers, we propose to add an additional step to the calculation of their GAFs to take into consideration these issuers' distance from the confidence interval. The present formula for an issuers' GAF, $GAF_{G,i} = GFR_{G,i} - \mu\{GFR_G\}$, would be

[50] In the 2020 Payment Notice final rule, we stated that we may consider alternative options for error rate adjustments, such as using multiple or smoothed confidence intervals for outlier identification and risk score adjustments. See 84 FR at 17507.

modified by replacing the $GFR_{G,i}$ with a decomposition of this value that uses the national weighted mean and national weighted standard deviation for the HCC failure rate group, as well as $z_{G,i}$, the z-score associated with the $GFR_{G,i}$,, where:

$$z_{G,i} = \frac{GFR_{G,i} - \mu\{GFR_G\}}{Sd\{GFR_G\}}$$

And therefore:

$$GFR_{G,i} = z_{G,i} * Sd\{GFR_G\} + \mu\{GFR_G\}$$

So:

$$GAF_{G,i} = \left[z_{G,i} * Sd\{GFR_G\} + \mu\{GFR_G\}\right] - \mu\{GFR_G\}$$

The z-score would then be discounted using the general formula: $disZ_{G,i,r} = a * z_{G,i} + b_r$, where $disZ_{G,i,r}$ is the confidence-level discounted z-score for that value of $z_{G,i}$ according to the parameters of the positive or negative sliding scale range, $r$. This $disZ_{G,i,r}$ value would replace the $z_{G,i}$ value in the $GAF_{G,i}$ formula to provide the value of the sliding scale adjustment for the positive or negative side of the confidence interval:

$$GAF_{G,i,r} = \left[disZ_{G,i,r} * Sd\{GFR_G\} + \mu\{GFR_G\}\right] - \mu\{GFR_G\}$$

In the calculation of $disZ_{G,i,r}$, the coefficient $a$ would be the slope of the linear adjustment, which shows the adjustment increase rate per unit increase of $GFR_{G,i}$, and $b_r$ is the intercept of the linear adjustment for either the negative or positive sliding scale range. The coefficients would be determined based on the standard deviation thresholds of the range selected for the application of the sliding scale adjustment. Specifically, coefficient $a$ would be defined as:

$$a = \frac{outerZ_r}{outerZ_r - innerZ_r}$$

Where:

- $a$ is the slope of the sliding scale adjustment

- *r* indicates whether the *GAF* is being calculated for a negative or positive outlier
- $outerZ_r$ is the greater magnitude z-score selected to define the edge of a given sliding scale range *r* (3.00 for positive outliers; and -3.00 for negative outliers)
- $innerZ_r$ is the lower magnitude z-score selected to define the edge of a given sliding scale range *r* (1.645 for positive outliers; and -1.645 for negative outliers)

The value of intercept $b_r$ would differ based on whether the sliding scale were being calculated for a positive or negative outlier and would be defined as:

$$b_r = outerZ_r - a * (outerZ_r) = outerZ_r * (1 - a)$$

In the absence of the constraints on negative failure rates described later in this proposed rule, the final formula for the group adjustment when an outlier issuer is subject to the sliding scale ($GAF_{G,i,r}$, above) could be simplified to:

$$GAF_{G,i,r} = disZ_{G,i,r} * Sd\{GFR_G\}$$

However, for the purposes of aligning formulas between the multiple proposals in this proposed rule, we feel that it is helpful to provide both the above expanded and simplified versions of the sliding scale $GAF_{G,i,r}$ formula in this section.

This sliding scale $GAF_{G,i,r}$ would be applied to the HCC coefficients in the applicable HCC failure rate group when calculating each enrollee with an HCCs' risk score adjustment factor for an issuer that had a failure rate with a z-score within the range of values selected for the sliding scale adjustment ($innerZ_r$ and $outerZ_r$). All other enrollee adjustment factors would be calculated using the current formula for the $GAF_{G,i}$. Using this linear sliding scale adjustment would provide a smoothing effect in the error rate calculation for issuers with failure rates just outside of the confidence interval of an HCC group.

To implement this proposed option, we would need to select the thresholds of the range (*inner$Z_r$ and outer$Z_r$)* to calculate and apply the sliding scale adjustment.[51] Commenters to the 2019 RADV White Paper supported a sliding scale option that would calculate and apply the sliding scale adjustment from +/- 1.96 to 3 standard deviations. This option would retain the confidence interval at 1.96 standard deviations under the current methodology, meaning that issuers within the 95 percent confidence interval would not have their respective risk scores adjusted. This option would also retain the full adjustment to the mean failure rate for issuers outside of the 99.7 percent confidence interval (beyond 3 standard deviations). While some of these stakeholders would prefer that the error rate be calculated to the edge of the confidence intervals for all outliers, rather than applying a sliding scale, some of these same commenters expressed support for this option because it would not increase the number of outliers compared to the current methodology, promoting stability for issuers. Specifically, this option would provide stability by maintaining the current thresholds used in the error rate calculation and without changing the number of issuers that would be impacted. While we recognize that this option would mitigate the payment cliff, we have concerns that it would weaken the HHS-RADV program by reducing its overall impact and the magnitude of HHS-RADV adjustments to the risk scores of outlier issuers.

Instead, in this proposed rule, we propose to calculate and apply a sliding scale adjustment between the 90 and 99.7 percent confidence interval bounds (from +/- 1.645 to 3

[51] In the 2019 RADV White Paper, we considered four different options on how to calculate and apply additional thresholds for the sliding scale adjustment to the error rate calculation. See section 4.4.4 and 4.4.5 of the 2019 RADV White Paper.

standard deviations). Under this proposal, the determination of outliers in HHS-RADV for each HCC grouping would no longer have a 95 percent confidence interval or 1.96 standard deviations, and would instead have a 90 percent confidence interval or 1.645 standard deviations. Specifically, this approach would adjust the upper and lower bounds of the confidence interval to be at 1.645 standard deviations, meaning that issuers outside of the 90 percent confidence interval would have their risk scores adjusted, instead of beginning adjustments for issuers at the 95 percent confidence interval under the current methodology. This would mean that more issuers would be considered outliers under this proposal than the current methodology.

Under this proposed approach, the above formulas would be implemented[52] as follows:

If $Freq\_EDGE_{G,i} \geq 30$, then:

If $z_{G,i} < -3.00$ or $z_{G,i} > 3.00$

Then $Flag_{G,i} = "outlier"$ and $GAF_{G,i} = GFR_{G,i} - \mu\{GFR_G\}$

Or if $-3 < z_{G,i} < -1.645$ or $3 > z_{G,i} > 1.645$

Then $Flag_{G,i} = "outlier"$ and $GAF_{G,i} = disZ_{G,i,r} * Sd\{GFR_G\}$

If $Freq\_EDGE_{G,i} < 30$ or if $-1.645 \leq z_{G,i} \leq 1.645$:

Then $Flag_{G,i} = "not\ outlier"$ and $GAF_{G,i} = 0$

Where $disZ_{G,i,r}$ is calculated using 3.00 (or -3.00, for negative outliers) as the value of *outerZr* and 1.645 (or -1.645, for negative outliers) as the value of *innerZr*.

[52] This calculation sequence is expressed here in a revised order compared to how the sequence is published in the 2021 Payment Notice (85 FR 29164 at 29196-29198). This change was made for simplicity to demonstrate how the current sequence would be combined with this proposed approach. The different display does not modify or otherwise change the amendments to the outlier identification process finalized in the 2021 Payment Notice.

This proposed approach would retain the current significant adjustment to the HCC group weighted mean for issuers beyond three standard deviations to ensure that the mitigation of the payment cliff for those issuers close to the confidence intervals does not impact situations where outlier issuers' failure rates are not close to the confidence intervals and a larger adjustment is warranted.

As discussed in the 2019 RADV White Paper, we tested a sliding scale adjustment between the 90 and 99 percent confidence interval bounds using 2017 HHS-RADV results.[53] We found that even though it would increase the number of outliers by including issuers whose failure rates fell between 1.645 and 1.96 standard deviations from the mean, it would lower the overall impact of HHS-RADV adjustments to transfers and result in the distribution of issuers' error rates moving closer to zero compared to the current methodology.[54] Therefore, this proposal preserves a strong incentive for issuers to submit accurate EDGE data that can be validated in HHS-RADV because it increases the range in which issuers can be flagged as outliers, while lowering the calculation of that adjustment amount for those outlier issuers close to the confidence intervals and maintaining a larger adjustment for those who are not close to the confidence intervals. For these reasons, we believe that this proposal for calculating and applying the sliding scale adjustment provides a balanced approach to addressing the payment cliff. We seek comment on this proposal, including the proposed calculation of the sliding scale adjustment and the thresholds used to calculate and apply it.

[53] See section 4.4.5 and Appendix C of the 2019 RADV White Paper.
[54] Ibid.

3. Negative Error Rate Issuers with Negative Failure Rates

HHS-RADV is intended to promote confidence and stability in the budget neutral HHS-operated risk adjustment program by ensuring the integrity and quality of data provided by issuers. HHS-RADV also serves to ensure that, consistent with the statute, charges are collected from issuers with lower-than-average actuarial risk and payments are made to issuers with higher-than-average actuarial risk. It uses a two-sided outlier identification approach because the long-standing intent of HHS-RADV has been to account for identified material risk differences between what issuers submitted to their EDGE servers and what was validated in medical records through HHS-RADV, regardless of the direction of those differences.[55] In addition, the two-sided adjustment policy penalizes issuers who validate HCCs in HHS-RADV at much lower rates than the national average and rewards issuers in HHS-RADV who validate HCCs in HHS-RADV at rates that are much higher than the national average, encouraging issuers to ensure that their EDGE-reported risk scores reflect the true actuarial risk of their enrollees. Positive and negative error rate outliers represent these two types of adjustments, respectively.

If an issuer is a positive error rate outlier, its risk score will be adjusted downward. Assuming no changes to risk scores for the other issuers in the same state market risk pool, this downward adjustment increases the issuer's charge or decreases its payment for the applicable benefit year, leading to a decrease in charges or an increase in payments for the other issuers in the state market risk pool. If an issuer is a negative error rate outlier, its risk score will be

[55] An exception to this approach was established, beginning with the 2018 benefit year of HHS-RADV, for exiting issuers who are negative error rate outliers. See 84 FR at 17503 – 17504.

adjusted upward. Assuming no changes to risk scores for the other issuers in the same state market risk pool, this upward adjustment reduces the issuer's charge or increases its payment for the applicable benefit year, leading to an increase in charges or a decrease in payments for the other issuers in the state market risk pool. The increase to risk score(s) for negative error rate outliers is consistent with the upward and downward risk score adjustments finalized as part of the original HHS-RADV methodology in the 2015 Payment Notice[56] and the HCC failure rate approach to error estimation finalized in the 2019 Payment Notice.[57] As noted above, some stakeholders have recommended HHS not adjust issuers' risk scores in the case of negative error rate issuers to limit the impact of these adjustments on issuers who are not outliers.

An issuer can be identified as a negative error rate outlier for a number of reasons. However, the current error rate methodology does not distinguish between low failure rates due to accurate data submission and failure rates that have been depressed through the presence of found HCCs (that is, HCCs in the audit data that were not present in the EDGE data). If a negative failure rate is due to a large number of found HCCs, it does not reflect accurate reporting through the EDGE server for risk adjustment. While we believe that any issuer with a negative failure rate is likely to review their internal processes to better capture missing HCCs in their future EDGE data submissions, we are proposing to refine the current error rate calculation

[56] For example, we stated that "the effect of an issuer's risk score error adjustment will depend upon its magnitude and direction compared to the average risk score error adjustment and direction for the entire market." See 79 FR 13743 at 13769.

[57] See 83 FR 16930 at 16962. The shorthand "positive error rate outlier" captures those issuers whose HCC coefficients are reduced as a result of being identified as an outlier, while "negative error rate outlier" captures those issuers whose HCC coefficients are increased as a result of being identified as an outlier.

to mitigate the impact of adjustments that result from negative error rate outliers whose low failure rates are driven by newly found HCCs rather than by high validation rates. We believe that a constraint in the GAF calculation in the current error rate calculation would mitigate potential incentives for issuers to use HHS-RADV to identify more HCCs than were reported to their EDGE servers. It also would mitigate the impact of HHS-RADV adjustments to transfers in the case of negative error rate issuers with negative failure rates and improve predictability.

Currently, an outlier issuer's error rate is calculated based on the difference between the weighted mean failure rate for the HCC group and the issuer's failure rate for that HCC grouping, which may be a negative failure rate. Beginning with 2019 benefit year HHS-RADV, we propose to adopt an approach that constrains negative error rate outlier issuers' error rate calculations in cases when an issuer's failure rate is negative. The proposed constraint would be to the GAF whereby the error rates of a negative error rate outlier issuer with a negative failure rate would be calculated as the difference between the weighted mean failure rate for the HCC grouping (if positive) and zero (0). This would be calculated by substituting the following ||double bars|| terms into the error rate calculation[58] process:

If $Freq\_EDGE_{G,i} \geq 30$, then:

If $GFR_{G,i} > UB_G$ or $GFR_{G,i} < LB_G$:

Then $Flag_{G,i} = "outlier"$ and $GAF_{G,i} = ||GFR_{G,i,constr} - \mu\{GFR_G\}_{constr}||$

[58] This calculation sequence is expressed here in a revised order compared to how the sequence is published in the 2021 Payment Notice (85 FR 29164 at 29196-29198). This change was made for simplicity when demonstrating how this sequence would be combined with this proposal. The different display does not modify or otherwise change the amendments to the outlier identification process finalized in the 2021 Payment Notice.

If $Freq\_EDGE_{G,i} < 30$ or if $GFR_{G,i} \leq UB_G$ and $GFR_{G,i} \geq LB_G$:

Then $Flag_{G,i} = "not\ outlier"$ and $GAF_{G,i} = 0$

Where:

$GFR_{G,i}$ is an issuer's failure rate for the HCC failure rate grouping

$||GFR_{G,i,constr}$ is an issuer's failure rate for the HCC failure rate grouping, constrained to 0 if is less than 0. Also expressed as:

$GFR_{G,i,constr} = max\{0, GFR_{G,i}\}||$

$\mu\{GFR_G\}$ is the weighted national mean failure rate for the HCC failure rate grouping

$||\mu\{GFR_G\}_{constr}$ is the weighted national mean failure rate for the HCC failure rate grouping, constrained to 0 if $\mu\{GFR_G\}$ is less than 0. Also expressed as:

$\mu\{GFR_G\}_{constr} = max\{0, \mu\{GFR_G\}\}||$

$UB_G$ and $LB_G$ are the upper and lower bounds of the HCC failure rate grouping confidence interval, respectively.

$Flag_{G,i}$ is the indicator if issuer *i*'s group failure rate for group G locates beyond a calculated threshold that we are using to classify issuers into "outliers" or "not outliers" for group *G*.

$GAF_{G,i}$ is the calculated adjustment amount to adjust issuer i's EDGE risk score components for all sampled HCCs in group *G*.

We would then compute total adjustments and error rates for each outlier issuer based on the weighted aggregates of the $GAF_{G,i.}$[59]

This approach would limit the financial impact that negative error rate outliers with negative failure rates would have on other issuers in the same state market risk pool, and would help provide stability to issuers in predicting the impact of HHS-RADV adjustments. For example, under the current error rate methodology using the 2017 benefit year HHS-RADV metrics, a negative outlier issuer with a -15 percent failure rate for the low HCC grouping would receive a GAF of the difference between -15 percent and the weighted mean for the low HCC grouping of 4.8 percent of -19.8 percent. However, under the proposal in this rulemaking to constrain the negative failure rates for negative outlier issuers to zero, the GAF in this example would be the difference between 0 percent and the weighted mean for the low HCC grouping of 4.8 percent, resulting in a -4.8 percent GAF.

If this proposal is finalized, the constrained values in the calculation of the GAF would only impact issuers with negative failure rates; therefore, issuers who have been extremely accurate in reporting their data to their EDGE server will not be affected. Issuers who report accurately to their EDGE servers are likely to have failure rates very close to zero, and may have negative error rates, but not negative failure rates. As such, these issuers would not have their GAF values constrained. In contrast, the issuers found to have negative failure rates, indicating that diagnosis data to their EDGE server was underreported for a particular benefit year, would

[59] See, for example, the 2018 Benefit Year Protocols: PPACA HHS Risk Adjustment Data Validation, Version 7.0 (June 24, 2019), available at: https://www.regtap.info/uploads/library/HRADV_2018Protocols_070319_5CR_070519.pdf.

have their GAF values constrained. As such, the proposed constraints on the GAF calculation will not apply or impact adjustments for issuers who are extremely accurate in reporting their diagnosis data to their EDGE servers.

We are proposing this option because it could be easily implemented under the current error rate methodology, would address stakeholders' concerns about the impact of adjustments due to negative error rate issuers with negative failure rates, and would reduce incentives that may exist for issuers to use HHS-RADV to identify more HCCs than existed in EDGE. We seek comment on this proposal.

a. Combining the HCC Grouping Constraint, Negative Failure Rate Constraint and the Sliding Scale Proposals

To help commenters understand the interaction of the above proposals to create Super HCCs for grouping purposes, apply a sliding scale option, and constrain negative failure rates for negative error rate outliers, this section outlines the complete proposed revised error rate calculation methodology formulas, integrating all the changes proposed to apply beginning with 2019 HHS-RADV in this proposed rule.

First, HHS would use the failure rates for Super HCCs to group each HCC into three HCC groupings (a high, medium, or low HCC failure rate grouping). Under the above proposed approach, Super HCCs would be defined as HCCs that have been aggregated such that HCCs that are in the same HCC coefficient estimation group are aggregated together and all other HCCs each compose an individual Super HCC. Using the Super HCCs, we would calculate the HCC failure rate as follows:

$$FR_c = 1 - \frac{freqIVA_c}{freqEDGE_c}$$

Where:

$c$ is the index of the $c^{th}$ Super HCC;

$freqEDGE_c$ is the frequency of a Super HCC $c$ occurring in EDGE data; that is, the sum of $freqEDGE_h$ for all HCCs that share an HCC coefficient estimation group in the adult risk adjustment models:

$$freqEDGE_c = \sum_h freqEDGE_{h,c}$$

When an HCC is not in an HCC coefficient estimation group in the adult risk adjustment models, the $freqEDGE_c$ for that HCC will be equivalent to $freqEDGE_h$;

$freqIVA_c$ is the frequency of a Super HCC $c$ occurring in IVA results (or SVA results, as applicable); that is, the sum of $freqIVA_h$ for all HCCs that share an HCC coefficient estimation group in the adult risk adjustment models:

$$freqIVA_c = \sum_h freqIVA_{h,c}$$

And;

$FR_c$ is the national overall (average) failure rate of Super HCC $c$ across all issuers.

Then, the failure rates for all Super HCCs, both those composed of a single HCC and those composed of the aggregate frequencies of HCCs that share an HCC coefficient estimation group in the adult models, would be grouped according to the current HHS-RADV failure rate grouping methodology. These HCC groupings would be determined by first ranking all Super HCC failure rates and then dividing the rankings into the three groupings weighted by total observations of that Super HCC across all issuers' IVA samples, assigning each Super HCC into a high, medium, or low HCC grouping. This process ensures that all HCCs in a Super HCC are grouped into the same HCC grouping in HHS-RADV.

Next, an issuer's HCC group failure rate would be calculated as follows:

$$GFR_{G,i} = 1 - \frac{freqIVA_{G,i}}{freqEDGE_{G,i}}$$

Where:

$freqEDGE_{G,i}$ is the number of occurrences of HCCs in group *G* that are recorded on EDGE for all enrollees sampled from issuer *i.*.

$freqIVA_{G,i}$ is the number of occurrences of HCCs in group *G* that are identified by the IVA audit (or SVA audit, as applicable) for all enrollees sampled from issuer *i.*

$GFR_{G,i}$ is issuer *i*'s group failure rate for the HCC group *G.*

HHS calculates the weighted mean failure rate and the standard deviation of each HCC group as:

$$\mu\{GFR_G\} = 1 - \frac{\sum_i freqIVA_{G,i}}{\sum_i freqEDGE_{G,i}}$$

$$Sd\{GFR_G\} = \sqrt{\frac{\sum_i \left(freqEDGE_{G,i} * \left(GFR_{G,i} - \mu\{GFR_G\}\right)^2\right)}{\sum_i freqEDGE_{G,i}}}$$

Where:

$\mu\{GFR_G\}$ is the weighted mean of $GFR_{G,i}$ of all issuers for the HCC group *G* weighted by all issuers' sample observations in each group.

$Sd\{GFR_G\}$ is the weighted standard deviation of $GFR_{G,i}$ of all issuers for the HCC group *G*.

Each issuer's HCC group failure rates would then be compared to the national metrics for each HCC grouping. If an issuer's failure rate for an HCC group falls outside of the two-tailed 90 percent confidence interval with a 1.645 standard deviation cutoff based on the weighted

mean failure rate for the HCC group, the failure rate for the issuer's HCCs in that group would be considered an outlier (if the issuer meets the minimum number of HCCs for the HCC group). Based on issuers' failure rates for each HCC group, outlier status would be determined for each issuer independently for each issuer's HCC failure rate group such that an issuer may be considered an outlier in one HCC failure rate group but not an outlier in another HCC failure rate group. Beginning with the 2019 benefit year, issuers will not be considered an outlier for an HCC group in which the issuer has fewer than 30 HCCs. If no issuers' HCC group failure rates in a state market risk pool materially deviate from the national mean of failure rates (that is, no issuers are outliers), HHS does not apply any adjustments to issuers' risk scores or to transfers in that state market risk pool.

Then, once the outlier issuers are determined, we would calculate the group adjustment factor taking into consideration the outlier issuer's distance from the confidence interval and limiting calculation of the group adjustment factor when the issuer has a negative failure rate. The formula[60] would apply as follows:

If $Freq\_EDGE_{G,i} \geq 30$, then:

If $z_{G,i} < -3.00$ or $z_{G,i} > 3.00$

Then $Flag_{G,i} = "outlier"$ and

$$GAF_{G,i} = \max\{0, GFR_{G,i}\} - \max\{0, \mu\{GFR_G\}\}$$

[60] This calculation sequence is expressed here in a revised order compared to how the sequence is published in the 2021 Payment Notice (85 FR 29164 at 29196-29198). This change was made for simplicity to demonstrate how this sequence would be combined with proposals in this proposed rule. The different display does not modify or otherwise change the amendments to the outlier identification process finalized in the 2021 Payment Notice.

Or if $-3 < z_{G,i} < -1.645$ or $3 > z_{G,i} > 1.645$

Then $Flag_{G,i} = "outlier"$ and

$$GAF_{G,i} = \max\{0, (disZ_{G,i,r} * Sd\{GFR_G\} + \mu\{GFR_G\})\} - \max\{0, \mu\{GFR_G\}\}$$

If $Freq\_EDGE_{G,i} < 30$ or if $-1.645 \leq z_{G,i} \leq 1.645$

Then $Flag_{G,i} = "not\ outlier"$ and $GAF_{G,i} = 0$

Where:

- *r* indicates whether the *GAF* is being calculated for a negative or positive outlier;
- *a* is the slope of the sliding scale adjustment, calculated as:

$$a = \frac{outerZ_r}{outerZ_r - innerZ_r}$$

With $outerZ_r$ defined as the greater magnitude z-score selected to define the edge of the sliding scale range *r* (3.00 for positive outliers; and -3.00 for negative outliers) and $innerZ_r$ defined as the lower magnitude z-score selected to define the edge of the range *r* (1.645 for positive outliers; and -1.645 for negative outliers);

- $b_r$ is the intercept of the sliding scale adjustment for a given sliding scale range *r*, calculated as:

$$b_r = outerZ_r - a * (outerZ_r) = outerZ_r * (1 - a)$$

- $disZ_{G,i,r}$ is the z-score of issuer *i*'s $GFR_{G,i,}$ for HCC failure rate group *G* discounted according to the sliding scale for range *r*, calculated as:

$$disZ_{G,i,r} = a * z_{G,i} + b_r$$

With $z_{G,i}$ defined as the z-score of *i* issuers' $GFR_{G,i}$:

$$z_{G,i} = \frac{GFR_{G,i} - \mu\{GFR_G\}}{Sd\{GFR_G\}}$$

- $GAF_{G,i}$ is the group adjustment factor for HCC failure rate group *G* for an issuer *i;*

- $Sd\{GFR_G\}$ is the weighted national standard deviation of all issuers' *GFR*s for HCC failure rate group *G;*
- $\mu\{GFR_G\}$ is the weighted national mean of all issuers' *GFR*s for HCC failure rate group *G.*

Once an outlier issuer's group adjustment factor is calculated, the enrollee adjustment would be calculated by applying the group adjustment factor to an enrollee's individual HCCs. For example, if an issuer has one enrollee with the HIV/AIDS HCC and the issuer's HCC group adjustment rate is 10 percent for the HCC group that contains the HIV/AIDS HCC, the enrollee's HIV/AIDS coefficient would be reduced by 10 percent. This reduction would be aggregated with any reductions to other HCCs for that enrollee to arrive at the overall enrollee adjustment factor. This value would be calculated according to the following formula for each sample enrollee in stratum 1 through 9:

$$Adjustment_{i,e} = \frac{\sum_h \left(RS_{h,G,i,e} * GAF_{G,i}\right)}{\sum_h \left(RS_{h,G,i,e}\right)}$$

Where:

$RS_{h,G,i,e}$ is the risk score component of a single HCC *h* (belonging to HCC group *G*) recorded on EDGE for enrollee *e* of issuer *i.*

$GAF_{G,i}$ is the group adjustment factor for HCC failure rate group *G* for an issuer *i;*

$Adjustment_{i,e}$ is the calculated adjustment amount to adjust enrollee *e* of issuer *i*'s EDGE risk scores.

The calculation of the enrollee adjustment factor only considers risk score factors related to the HCCs and ignores any other risk score factors (such as demographic factors and RXC factors). Furthermore, because this formula is concerned exclusively with EDGE HCCs, HCCs

newly identified by the IVA (or SVA as applicable) would not contribute to enrollee risk score adjustments for that enrollee and adjusted enrollee risk scores are only computed for sampled enrollees with HCCs in strata 1 through 9.

Next, for each sampled enrollee with HCCs, HHS would calculate the total adjusted enrollee risk score as:

$$AdjRS_{i,e} = EdgeRS_{i,e} * (1 - Adjustment_{i,e})$$

Where:

$EdgeRS_{i,e}$ is the risk score as recorded on the EDGE server of enrollee *e* of issuer *i*.

$AdjRS_{i,e}$ is the amended risk score for sampled enrollee *e* of issuer *i*.

$Adjustment_{i,e}$ is the adjustment factor by which we estimate whether the EDGE risk score exceeds or falls short of the initial or second validation audit projected total risk score for sampled enrollee *e* of issuer *i.*

The calculation of the sample enrollee's adjusted risk score includes all EDGE server components for sample enrollees in strata 1 through 9.

After calculating the outlier issuers' sample enrollees with HCCs' adjusted EDGE risk scores, HHS would calculate an outlier issuer's error rate by extrapolating the difference between the amended risk score and EDGE risk score for all enrollees (stratum 1 through 10) in the sample. The extrapolation formula would be weighted by determining the ratio of an enrollee's stratum size in the issuer's population to the number of sample enrollees in the same stratum as the enrollee. Sample enrollees with no HCCs would be included in the extrapolation of the error rate for outlier issuers with the EDGE risk score unchanged for these sample enrollees. The formulas to compute the error rate using the stratum-weighted risk score before and after the adjustment would be:

$$ErrorRate_i = 1 - \frac{\sum_e (w_{i,e} * AdjRS_{i,e})}{\sum_e (w_{i,e} * EdgeRS_{i,e})}$$

Where:

$$w_{i,e} = \frac{stratum\ size\ in\ population}{number\ of sample\ enrollees\ of\ the\ stratum}$$

Consistent with 45 CFR 153.350(b), HHS then would apply the outlier issuer's error rate to adjust that issuer's applicable benefit year's plan liability risk score.[61] This risk score change, which also would impact the state market average risk score, would then be used to adjust the applicable benefit year's risk adjustment transfers for the applicable state market risk pool.[62] Due to the budget-neutral nature of the HHS-operated program, adjustments to one issuer's risk scores and risk adjustment transfers based on HHS-RADV findings affects other issuers in the state market risk pool (including those who were not identified as outliers) because the state market average risk score changes to reflect the outlier issuer's change in its plan liability risk score. This also means that issuers that are exempt from HHS-RADV for a given benefit year will have their risk adjustment transfers adjusted based on other issuers' HHS-RADV results if any issuers in the applicable state market risk pool are identified as outliers. We seek comments on our modified error rate calculation methodology proposed to be applicable starting for the 2019 benefit year of HHS-RADV.

[61] Exiting outlier issuer risk score error rates are currently applied to the plan liability risk scores and risk adjustment transfer amounts for the benefit year being audited if they are a positive error rate outlier. For all other outlier issuers, risk score error rates are currently applied to the plan liability risk scores and risk adjustment transfer amounts for the current transfer year. The exiting issuer exception is discussed in Section II.B.

[62] See 45 CFR 153.350(c).

In drafting this proposed rule, as requested by commenters on the 2019 RADV White Paper, we estimated the combined impact of applying the proposed sliding scale adjustment, the proposed negative failure rate constraint and the proposed Super HCC aggregation using 2017 benefit year HHS-RADV results. Table 5 provides a comparison of the estimated change in error rates between the current methodology for sorting HCCs for HHS-RADV grouping and the proposed Super HCC aggregation for sorting of HCCs for HHS-RADV grouping, the proposed negative failure rate constraint and the proposed sliding scale option in this proposed rule. In addition, in response to comments on the 2019 RADV White Paper that supported the adoption of a sliding scale adjustment from +/- 1.96 to 3 standard deviations, Table 5 also includes information on the estimated change(s) if option 1 from the 2019 RADV White Paper was adopted as the sliding scale adjustment.

As shown in Table 5, we also found through testing the 2017 benefit year HHS-RADV results that, although the proposed sliding scale adjustment (adjusting from +/-1.645 to 3 standard deviations) increases the number of outliers, the mean error rates among positive outliers under this proposal are smaller than the mean error rates among positive outliers for the 2019 RADV White Paper sliding scale option 1 (adjusting from +/- 1.96 to 3 standard deviations), even when tested in combination with the proposed negative failure rate constraint and/or the current and proposed sorting methodologies. This suggests that the proposed sliding scale option would result in reduced HHS-RADV adjustments to risk adjustment transfers relative to both the current methodology and the 2019 RADV White Paper sliding scale option 1, and reflects the smoother transition between a GAF of zero and a full-value GAF that is provided by the proposed sliding scale option when compared to 2019 RADV White Paper sliding scale option 1.

**Table 5: A Comparison of HHS-RADV Error Rate (ER) Estimated Changes Based on 2017 Benefit Year[63] HHS-RADV Data**

| Scenario | Current Sorting Method | | Super HCCs using HCC Coefficient Estimation Groups | |
|---|---|---|---|---|
| | Mean Neg ER | Mean Pos ER | Mean Neg ER | Mean Pos ER |
| Sorting Method Only | -5.68% | 9.96% | -5.98% | 9.91% |
| Sorting Method with Proposed Negative Constraint | -3.11% | 9.96% | -3.38% | 9.91% |
| Sorting Method with Proposed Sliding Scale Option[64] | -2.27% | 5.28% | -2.49% | 5.32% |
| Sorting Method, Proposed Sliding Scale Option & Proposed Negative Constraint | -1.50% | 5.28% | -1.66% | 5.32% |
| Sorting Method with 2019 RADV White Paper Sliding Scale Option 1[65] | -2.16% | 6.46% | -2.48% | 6.51% |
| Sorting Method with 2019 RADV White Paper Sliding Scale Option 1 & Proposed Negative Constraint | -1.12% | 6.46% | -1.26% | 6.51% |

## B. Application of HHS-RADV Results

In the 2014 Payment Notice, HHS finalized a prospective approach for making adjustments to risk adjustment transfers based on findings from the HHS-RADV process.[66]

[63] These estimates include the exclusion from outlier status of issuers with fewer than 30 HCCs in an HCC group, consistent with the policy finalized in the 2021 Payment Notice (85 FR 29164), which was not in effect for 2017 Benefit Year HHS-RADV. We included the fewer than 30 HCC exclusion from outlier status in these estimates to provide a sense of the impact of the proposed changes when compared to the methodology presently in effect for 2019 benefit year HHS-RADV and beyond.

[64] The Proposed Sliding Scale Option outlined in Section II.A.2. of this rule would create a sliding scale adjustment from +/-1.645 to 3 standard deviations.

[65] The 2019 RADV White Paper Sliding Scale Option 1 would create a sliding scale adjustment from +/- 1.96 to 3 standard deviations.

Specifically, we finalized using an issuer's HHS-RADV error rates from the prior year to adjust the issuer's average risk score in the current benefit year. As such, we used the 2017 benefit year HHS-RADV results to adjust 2018 benefit year risk adjustment plan liability risk scores for non-exiting issuers, resulting in adjustments to 2018 benefit year risk adjustment transfer amounts.[67,68]

When we finalized the prospective HHS-RADV results application policy in the 2014 Payment Notice, we did not anticipate the extent of the changes that could occur in the risk profile of enrollees or market participation in the individual and small group markets from benefit year to benefit year. As a result of experience with these changes over the early years of the program, and in light of the changes finalized in the 2020 Payment Notice to the timeline for the reporting, collection, and disbursement of risk adjustment transfer adjustments for HHS-RADV[69] and the changes to the risk adjustment holdback policy,[70] both of which will lead

---

[66] See 78 FR 15409 at 15438.
[67] See the Summary Report of 2017 Benefit Year HHS-RADV Adjustments to Risk Adjustment Transfers released on August 1, 2019, available at: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/BY2017-HHSRADV-Adjustments-to-RA-Transfers-Summary-Report.pdf.
[68] In the 2019 Payment Notice, we adopted an exception to the prospective application of HHS-RADV results for exiting issuers, whereby risk score error rates for outlier exiting issuers are applied to the plan liability risk scores and transfer amounts for the benefit year being audited. Therefore, for exiting issuers, we used the 2017 benefit year's HHS-RADV results to adjust 2017 benefit year risk adjustment plan liability risk scores, resulting in adjustments to 2017 benefit year risk adjustment transfer amounts. See 83 FR at 16965 through 16966.
[69] See 84 FR at 17504 through 17508.
[70] See the Change to Risk Adjustment Holdback Policy for the 2018 Benefit Year and Beyond Bulletin (May 31, 2019) (May 2019 Holdback Guidance), available at: https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Change-to-Risk-Adjustment-Holdback-Policy-for-the-2018-Benefit-Year-and-Beyond.pdf.

to reopening of prior year risk adjustment transfers, we are now proposing changes to this prospective approach for non-exiting issuers.

Starting with the 2021 benefit year of HHS-RADV, we propose applying HHS-RADV results to the benefit year being audited for all issuers. This proposal is intended to address stakeholder concerns about maintaining actuarial soundness in the application of an issuer's HHS-RADV error rate if an issuer's risk profile, enrollment, or market participation changes substantially from benefit year to benefit year. This proposed change has the potential to provide more stability for issuers of risk adjustment covered plans and help them better predict the impact of HHS-RADV results. It would also prevent situations where an issuer who newly enters a state market risk pool is subject to HHS-RADV adjustments from the prior benefit year for which they did not participate. We seek comment on this proposal.

If we finalize and implement the policy to adjust the benefit year being audited beginning with the 2021 benefit year HHS-RADV, we would need to adopt transitional measures to move from the current prospective approach to one that applies the HHS-RADV results to the benefit year being audited. More specifically, 2021 benefit year risk adjustment plan liability risk scores and transfers would need to be adjusted first to reflect 2020 benefit year HHS-RADV results, and adjusted again based on 2021 benefit year HHS-RADV results. For the 2022 benefit year of HHS-RADV and beyond, risk adjustment plan liability risk scores and transfers would only be adjusted once based on the same benefit year's HHS-RADV results (that is, 2022 benefit year

HHS-RADV results would adjust 2022 benefit year risk adjustment plan liability risk scores and transfers).[71]

In order to effectuate this transition, we considered and are proposing an "average error rate approach," as set forth in the 2019 RADV White Paper, under which HHS would calculate an average value for the 2021 and 2020 benefit years' HHS-RADV error rates and apply this average error rate to 2021 risk adjustment plan liability risk scores and transfers. This approach would result in one final HHS-RADV adjustment to 2021 benefit year risk adjustment plan liability risk scores and transfers, reflecting the average value for the 2021 and 2020 benefit years' HHS-RADV error rates. The adjustments to transfers would be collected and paid in accordance with the 2021 benefit year HHS-RADV timeline, in 2025.[72]

However, in an effort to be consistent with our current risk score error rate application and calculation and ensure that both years of HHS-RADV results are taken into consideration in calculating risk adjustment plan liability risk scores, we also propose as an alternative transition strategy from the prospective application of HHS-RADV results to a concurrent application approach the "combined plan liability risk score option," also set forth in the 2019 RADV White Paper. Under the combined plan liability risk score option, we would apply 2020 benefit year HHS-RADV risk score adjustments to 2021 benefit year plan liability risk scores, and then apply 2021 benefit year HHS-RADV risk score adjustments to the adjusted 2021 plan liability risk

[71] As discussed in the May 2019 Holdback Guidance, a successful HHS-RADV appeal may require additional adjustments to transfers for the applicable benefit year in the impacted state market risk pool.
[72] For a general description of the current timeline for reporting, collection, and disbursement of HHS-RADV adjustments to transfers, see 84 FR at 17506 through 17507.

scores. We would then use the final adjusted plan liability risk scores (reflecting both the 2020 and 2021 HHS-RADV adjustments to risk scores) to adjust 2021 benefit year transfers. Under this proposal, HHS would calculate risk score adjustments for 2020 and 2021 benefit year HHS-RADV sequentially and incorporate 2020 and 2021 benefit year HHS-RADV results in one final adjustment amount to 2021 benefit year transfers that would be collected and paid in accordance with the 2021 benefit year HHS-RADV timeline, in 2025. We seek comment on both of these approaches to transition from the current prospective approach to one that applies the HHS-RADV results to the benefit year being audited.

Additionally, the transition to a policy to apply HHS-RADV results to the benefit year being audited would remove the need to continue the current policy on issuers entering sole issuer markets that was finalized in the 2020 Payment Notice.[73] As finalized in the 2020 Payment Notice, new issuer(s) that enter a new market or a previously sole issuer market have their risk adjustment transfers in the current benefit year adjusted if there was an outlier issuer in the applicable state market risk pool in the prior benefit year's HHS-RADV.[74] If the proposal to apply HHS-RADV results to the benefit year being audited for all issuers is finalized, new issuers, including new issuers in previously sole issuer markets, would no longer be prospectively impacted by HHS-RADV results from a previous benefit year; rather, the new issuer would only have their current benefit year risk scores (and subsequently, risk adjustment transfers) impacted. The exception would be for the proposed transition benefit years, 2020 and

[73] 84 FR at 17504.
[74] Ibid.

2021. If a new issuer enters a market in 2021, its risk adjustment plan liability risk score and transfers could be impacted by the new issuer's own 2021 HHS-RADV results and the combined 2020 and 2021 HHS-RADV results of other issuers in the same state market risk pool(s). In addition, since the current prospective approach would continue to apply to the 2019 benefit year HHS-RADV, if a new issuer enters a sole issuer market in 2020, this new issuer would see its 2020 risk adjustment plan liability risk scores and transfers impacted if there was an outlier issuer as a result of 2019 benefit year HHS-RADV in the applicable state market risk pool.

We solicit comment on all of these proposals. In addition, in light of the postponement of the 2019 HHS-RADV process as part of the Administration's efforts to combat COVID-19,[75] we are additionally seeking comment on an alternative timeline for the proposed transition from the prospective application of HHS-RADV results for non-exiting issuers.

Under this alternative timeline, we would apply HHS-RADV results to the benefit year being audited for all issuers starting with the 2020 benefit year of HHS-RADV, rather than the 2021 benefit year. If we finalize and implement either of the above transition options using the alterative timeline, 2020 benefit year risk adjustment plan liability risk scores and transfers would need to be adjusted twice – first to reflect 2019 benefit year HHS-RADV results and again based on 2020 benefit year HHS-RADV results.[76] To accomplish this, we would either

[75] Available at https://www.cms.gov/files/document/2019-HHS-RADV-Postponement-Memo.pdf. As discussed in the memo, our intention is to provide guidance by August 2020 on the updated timeline for 2019 benefit year HHS-RADV activities that we plan to begin in 2021.

[76] If no changes are made to the timeline for 2020 benefit year HHS-RADV activities, they would begin with the release of enrollee samples in late May 2021. Given the postponement of 2019 benefit year HHS-RADV activities in response to the COVID-19 pandemic, it is possible HHS-RADV activities for the 2019 and 2020 benefit years would be conducted at the same time.

(1) implement the "combined plan liability risk score option," whereby we would apply 2019 benefit year HHS-RADV risk score adjustments to 2020 benefit year plan liability risk scores, and then apply 2020 benefit year HHS-RADV risk score adjustments to the already adjusted 2020 plan liability risk scores, or (2) implement the "average error rate approach," whereby we would calculate an average value for the 2019 and 2020 benefit years' HHS-RADV error rates and apply the averaged error rate to 2020 benefit year plan liability risk scores. We would then use the final adjusted plan liability risk scores from either of these approaches to adjust 2020 benefit year transfers. The adjustments to transfers would be collected and paid in accordance with the 2020 benefit year HHS-RADV timeline, in 2024. We also seek comment on whether, if we finalize and implement either of the above transition options using the alterative timeline, we should also pilot RXCs for the 2020 benefit year HHS-RADV to increase consistency between the operations of 2019 and 2020 HHS-RADV. We solicit comment on all of these proposals.

## III. Collection of Information Requirements

This document does not impose information collection requirements, that is, reporting, recordkeeping, or third-party disclosure requirements. Consequently, there is no need for review by the Office of Management and Budget under the authority of the Paperwork Reduction Act of 1995 (44 U.S.C. 3501 et seq.).

Under this proposed rule, we propose to amend the calculation of error rates to modify the sorting methodology for HCCs that share an HCC coefficient estimation group in the adult risk adjustment models; to amend the error rate calculation for cases where outlier issuers are near the confidence intervals; to constrain the error rate calculation for issuers with negative failure rates; and to transition to the application of HHS-RADV results to the benefit year being audited. These proposed changes are methodological changes to the error estimation

methodology used in calculating error rates and changes to the application of HHS-RADV results to risk scores and transfers. Since HHS calculates error rates and applies HHS-RADV results to risk scores and transfers, we do not estimate a burden change on issuers to conduct and complete HHS-RADV in states where HHS operates the risk adjustment program for a given benefit year. [77]

## IV. Response to Comments

Because of the large number of public comments, we normally receive on **Federal Register** documents, we are not able to acknowledge or respond to them individually. We will consider all comments we receive by the date and time specified in the "DATES" section of this proposed rule, and, when we proceed with a subsequent document, we will respond to the comments in the preamble to that document.

## V. Regulatory Impact Statement

### A. Statement of Need

This rule proposes standards related to the HHS-RADV program, including certain refinements to the calculation of error rates and a transition from the prospective application of HHS-RADV results. The Premium Stabilization Rule and other rulemakings noted above provided detail on the implementation of the HHS-RADV program.

[77] Since the 2017 benefit year, HHS has been responsible for operating risk adjustment in all 50 states and the District of Columbia.

B. Overall Impact

We have examined the impact of this rule as required by Executive Order 12866 on Regulatory Planning and Review (September 30, 1993), Executive Order 13563 on Improving Regulation and Regulatory Review (January 18, 2011), the Regulatory Flexibility Act (RFA) (September 19, 1980, Pub. L. 96-354), section 1102(b) of the Social Security Act (the Act), section 202 of the Unfunded Mandates Reform Act of 1995 (March 22, 1995; Pub. L. 104-4), Executive Order 13132 on Federalism (August 4, 1999), the Congressional Review Act (5 U.S.C. 804(2)), and Executive Order 13771 on Reducing Regulation and Controlling Regulatory Costs (January 30, 2017).

Executive Orders 12866 and 13563 direct agencies to assess all costs and benefits of available regulatory alternatives and, if regulation is necessary, to select regulatory approaches that maximize net benefits (including potential economic, environmental, public health and safety effects, distributive impacts, and equity). A Regulatory Impact Analysis (RIA) must be prepared for major rules with economically significant effects ($100 million or more in any 1 year). This rule does not reach the economic significance threshold, and thus is not considered a major rule. For the same reason, it is not a major rule under the Congressional Review Act.

C. Regulatory Alternatives Considered

In developing the policies contained in this proposed rule, we considered numerous alternatives to the presented proposals. Below we discuss the key regulatory alternatives considered.

We considered an alternative approach to the proposed sorting of all HCCs that share an HCC coefficient estimation group in the adult models into the same “Super HCC” for HHS-RADV HCC grouping purposes. This alternative approach would have combined all HCCs

in the same hierarchy into the same Super HCC for HHS-RADV HCC grouping purposes even if those HCCs had different coefficients in the risk adjustment models. While we did analyze this option, we were concerned that it would not account for risk differences within the HCC hierarchies, and that the proposed approach that focuses on HCCs with the same risk scores in the adult models would better ensure that HHS-RADV results account for risk differences within HCC hierarchies. Additionally, by forcing all HCCs that share a hierarchy into the same HHS-RADV failure rate grouping regardless of whether they have different coefficients, we would not only diminish our ability to allow for differences among various diseases within an HCC hierarchy but would also reduce our ability to recognize differences in the difficulty of providing medical documentation for them.[78]

We considered several other options for addressing the payment cliff effect besides the specific sliding scale approach that we proposed. One option was returning to the original methodology finalized in the 2015 Payment Notice, which would have adjusted almost all issuers' risk scores for every error identified as a result of HHS-RADV.[79] The adjustments under the original methodology would have used the issuer's corrected average risk score to compute an adjustment factor, which would have been based on the ratio between the corrected and original average risk scores. However, our analysis indicated that the original methodology generally resulted in a more severe payment cliff effect, since the majority of outlier issuers had

[78] See 83 FR 16961 and 16965.
[79] See 79 FR 13755-13770.

their original failure rates applied without the benefit of subtracting the weighted mean difference.[80]

The second option we considered was to modify the error rate calculation by calculating the issuer's GAF using the HCC group confidence interval rather than the distance to the weighted HCC group mean. As described in the 2019 RADV White Paper and in previous rulemaking,[81] we have concerns that this option would result in under-adjustments based on HHS-RADV results for issuers farthest from the confidence intervals. Thus, although this option could address the payment cliff effect for issuers just outside of the confidence interval, it also could create the unintended consequence of mitigating the payment impact for situations where issuers are not close to the confidence intervals, potentially reducing incentives for issuers to submit accurate risk adjustment data to their EDGE servers.

An additional option suggested by some stakeholders that could address, at least in part, the payment cliff effect that we considered would be to modify the current two-sided approach to HHS-RADV and only adjust issuers who are positive error rate outliers. However, moving to a one-sided outlier identification methodology would not have addressed the payment cliff effect because it would still exist on the positive error rate side of the methodology.[82] In addition, the two-sided outlier identification, and the resulting adjustments to outlier issuer risk scores that

[80] See the 2019 RADV White Paper at pages 78-79 and Appendix B.
[81] See 84 FR 17507 – 17508. See also the 2019 RADV White Paper at page 80.
[82] It is important to note the purpose of HHS-RADV approach is fundamentally different from the Medicare Advantage risk adjustment data validation (MA-RADV) approach. MA-RADV only adjusts for positive error rate outliers, as the program's intent is to recoup Federal funding that was the result of improper payments under the Medicare Part C program.

have significantly better-than-average or poorer-than-average data validation results, ensures that HHS-RADV adjusts for identified, material risk differences between what issuers submitted to their EDGE servers and what was validated by the issuers' medical records. The two-sided outlier identification approach ensures that an issuer who is coding well is able to recoup funds that might have been lost through risk adjustment because its competitors are coding badly.

We also considered various other options for the thresholds under the sliding scale option that we are proposing to address the payment cliff effect. For example, we considered as an alternative the adoption of a sliding scale option that would adjust outlier issuers' error rates on a sliding scale between the 95 and 99 percent confidence interval bounds (from +/- 1.96 to 3 standard deviations). This alternative sliding scale option would retain the current methodology's confidence interval at 1.96 standard deviations, the full adjustment to the mean failure rate for issuers outside of the 99 percent confidence interval (beyond three standard deviations), and the current significant adjustment to the HCC group weighted mean after three standard deviations. In comments on the 2019 RADV White Paper, stakeholders expressed support for this sliding-scale option because it addressed the payment cliff issue without increasing the number of issuers identified as outliers. However, while we recognize that this alternative also would address the payment cliff effect, we are concerned it would not provide the same balanced approach as the proposed sliding scale option and would instead weaken the HHS-RADV program by reducing its overall impact and the magnitude of HHS-RADV adjustments to outlier issuer's risk scores.

When developing a process for implementing the transition from the prospective application of HHS-RADV results to a concurrent application approach, we considered three options for the transition year. In previous sections of the proposed rule, we described two of

those options. The third option is the "RA transfer option." The RA transfer option would separately calculate 2020 benefit year HHS-RADV adjustments to 2021 benefit year transfers and 2021 benefit year HHS-RADV adjustments to 2021 benefit year transfers.[83] Under this option, we would then calculate the difference between each of these values and the unadjusted 2021 benefit year transfers before any HHS-RADV adjustments were applied, and add these differences together to arrive at the total HHS-RADV adjustment that would be applied to the 2021 benefit year transfers. That is, HHS would separately calculate adjustments for the 2020 and 2021 benefit year HHS-RADV results and incorporate 2020 and 2021 benefit year HHS-RADV results in one final adjustment to 2021 benefit year transfers that would be collected and paid in accordance with the 2021 benefit year HHS-RADV timeline, in 2025.[84] However, we believe this alternative is not as consistent with our current risk score error rate application and calculation as the combined plan liability risk score option, or as simple as the average error rate approach discussed above.

## VI. Regulatory Flexibility Act

The Regulatory Flexibility Act (5 U.S.C. 601, et seq.) (RFA), requires agencies to prepare an initial regulatory flexibility analysis to describe the impact of a proposed rule on small entities, unless the head of the agency can certify that the rule will not have a significant economic impact on a substantial number of small entities. The RFA generally defines a "small entity" as (1) a proprietary firm meeting the size standards of the Small Business Administration

[83] See section 5.2 of the 2019 RADV White Paper.
[84] For a general description of the current timeline for publication, collection, and distribution of HHS-RADV adjustments to transfers, see 84 FR at 17506 –17507.

(SBA), (2) a not-for-profit organization that is not dominant in its field, or (3) a small government jurisdiction with a population of less than 50,000. States and individuals are not included in the definition of "small entity." HHS uses a change in revenues of more than 3 to 5 percent as its measure of significant economic impact on a substantial number of small entities.

In this proposed rule, we propose standards for the HHS-RADV program. This program is generally intended to ensure the integrity of the HHS-operated risk adjustment program, which stabilizes premiums and reduces the incentives for issuers to avoid higher-risk enrollees. Because we believe that insurance firms offering comprehensive health insurance policies generally exceed the size thresholds for "small entities" established by the SBA, we do not believe that an initial regulatory flexibility analysis is required for such firms.

We believe that health insurance issuers would be classified under the North American Industry Classification System code 524114 (Direct Health and Medical Insurance Carriers). According to SBA size standards, entities with average annual receipts of $41.5 million or less would be considered small entities for these North American Industry Classification System codes. Issuers could possibly be classified in 621491 (HMO Medical Centers) and, if this is the case, the SBA size standard would be $35.0 million or less.[85] We believe that few, if any, insurance companies underwriting comprehensive health insurance policies (in contrast, for example, to travel insurance policies or dental discount policies) fall below these size thresholds. Based on data from MLR annual report[86] submissions for the 2017 MLR reporting year,

[85] https://www.sba.gov/document/support--table-size-standards.
[86] Available at https://www.cms.gov/CCIIO/Resources/Data-Resources/mlr.html.

approximately 90 out of 500 issuers of health insurance coverage nationwide had total premium revenue of $41.5 million or less. This estimate may overstate the actual number of small health insurance companies that may be affected, since over 72 percent of these small companies belong to larger holding groups, and many, if not all, of these small companies are likely to have non-health lines of business that will result in their revenues exceeding $41.5 million.

In addition, section 1102(b) of the Act requires us to prepare an RIA if a rule may have a significant impact on the operations of a substantial number of small rural hospitals. This analysis must conform to the provisions of section 603 of the RFA. For purposes of section 1102(b) of the Act, we define a small rural hospital as a hospital that is located outside of a metropolitan statistical area and has fewer than 100 beds. This proposed rule would not affect small rural hospitals. Therefore, the Secretary has determined that this proposed rule will not have a significant impact on the operations of a substantial number of small rural hospitals.

**VII. Unfunded Mandates**

Section 202 of the Unfunded Mandates Reform Act of 1995 (UMRA) requires that agencies assess anticipated costs and benefits and take certain other actions before issuing a proposed rule that includes any Federal mandate that may result in expenditures in any 1 year by state, local, or Tribal governments, in the aggregate, or by the private sector, of $100 million in 1995 dollars, updated annually for inflation. In 2019, that threshold is approximately $154 million. Although we have not been able to quantify all costs, we expect the combined impact on state, local, or Tribal governments and the private sector to be below the threshold.

**VIII. Federalism**

Executive Order 13132 establishes certain requirements that an agency must meet when it issues a proposed rule that imposes substantial direct costs on state and local governments, preempts state law, or otherwise has federalism implications.

In compliance with the requirement of Executive Order 13132 that agencies examine closely any policies that may have federalism implications or limit the policymaking discretion of the states, we have engaged in efforts to consult with and work cooperatively with affected states, including participating in conference calls with and attending conferences of the National Association of Insurance Commissioners, and consulting with state insurance officials on an individual basis.

While developing this proposed rule, we attempted to balance the states' interests in regulating health insurance issuers with the need to ensure market stability. By doing so, it is our view that we have complied with the requirements of Executive Order 13132.

Because states have flexibility in designing their Exchange and Exchange-related programs, state decisions will ultimately influence both administrative expenses and overall premiums. States are not required to establish an Exchange or risk adjustment program. HHS operates risk adjustment on behalf of any state that does not elect to do so. Beginning with the 2017 benefit year, HHS has operated risk adjustment for all 50 states and the District of Columbia.

In our view, while this proposed rule would not impose substantial direct requirement costs on state and local governments, it has federalism implications due to direct effects on the distribution of power and responsibilities among the state and Federal Governments relating to determining standards about health insurance that is offered in the individual and small group markets.

## IX. Reducing Regulation and Controlling Regulatory Costs

Executive Order 13771 requires that the costs associated with significant new regulations ‘‘to the extent permitted by law, be offset by the elimination of existing costs associated with at least two prior regulations.’’ This proposed rule is not subject to the requirements of Executive Order 13771 because it is expected to result in no more than de minimis costs.

## X. Conclusion

In accordance with the provisions of Executive Order 12866, this regulation was reviewed by the Office of Management and Budget.

**Dated:** February 19, 2020.

_____

**Seema Verma,**

Administrator,

Centers for Medicare & Medicaid Services.

**Dated:** May 20, 2020.

_____

**Alex M. Azar II,**

Secretary,

Department of Health and Human Services.

[FR Doc. 2020-11703 Filed: 5/29/2020 4:15 pm; Publication Date: 6/2/2020]