§ 1.6011–8 Requirement of income tax return for taxpayers who claim the premium tax credit under section 36B.

(a) *Requirement of return.* Except as otherwise provided in this paragraph (a), a taxpayer who receives the benefit of advance payments of the premium tax credit (advance credit payments) under section 36B must file an income tax return for that taxable year on or before the due date for the return (including extensions of time for filing) and reconcile the advance credit payments. However, if advance credit payments are made for coverage of an individual who is not included in any taxpayer's family, as defined in § 1.36B–1(d), the taxpayer who attested to the Exchange to the intention to include such individual in the taxpayer's family as part of the advance credit payment eligibility determination for coverage of the individual must file a tax return and reconcile the advance credit payments.

(b) *Applicability dates*—(1) *In general.* Except as provided in paragraph (b)(2) of this section, paragraph (a) of this section applies for taxable years ending on or after December 31, 2020.

(2) *Prior periods.* Paragraph (a) of this section as contained in 26 CFR part 1 edition revised as of April 1, 2016, applies to taxable years ending after December 31, 2013, and beginning before January 1, 2017. Paragraph (a) of this section as contained in 26 CFR part 1 edition revised as of April 1, 2020, applies to taxable years beginning after December 31, 2016, and ending before December 31, 2020.

Sunita Lough,

Deputy Commissioner for Services and Enforcement.

Approved: September 4, 2020.

David J. Kautter,

Assistant Secretary of the Treasury (Tax Policy).

[FR Doc. 2020–26200 Filed 11–27–20; 11:15 am]

BILLING CODE 4830–01–P

DEPARTMENT OF JUSTICE

Office of the Attorney General

28 CFR Part 26

[Docket Number OAG 171; AG Order No. 4911–2020]

RIN 1105–AB63

Manner of Federal Executions

Correction

In rule document 2020–25867 beginning on page 75846 in the issue of Friday, November 27, 2020, make the following correction:

On page 75846, in the third column, in the last line, ''December 24, 2020'' should read ''December 28, 2020.''

[FR Doc. C1–2020–25867 Filed 11–30–20; 8:45 am]

BILLING CODE 1301–00–D

DEPARTMENT OF HEALTH AND HUMAN SERVICES

Centers for Medicare & Medicaid Services

45 CFR Part 153

[CMS–9913–F]

RIN 0938–AU23

Amendments to the HHS-Operated Risk Adjustment Data Validation (HHS–RADV) 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: Final rule.

SUMMARY: This final rule adopts certain changes to the risk adjustment data validation error estimation methodology beginning with the 2019 benefit year for states where the Department of Health and Human Services (HHS) operates the risk adjustment program. This rule is finalizing 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 finalizes a change to the benefit year to which HHS–RADV adjustments to risk scores and risk adjustment transfers would be applied beginning with the 2020 benefit year of HHS–RADV. These policies seek to further the integrity of HHS–RADV, address stakeholder feedback, promote fairness, and improve the predictability of HHS–RADV adjustments.

DATES: These regulations are effective on December 31, 2020.

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:

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 final 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 of the Department of Health and Human Services (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 through 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 the Department of Health and Human Services (HHS). 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 scores 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.[4]

Pursuant to 45 CFR 153.350, HHS performs HHS–RADV to validate the accuracy of data submitted by issuers

[1] 42 U.S.C. 18063.

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

[3] 42 U.S.C. 18063(b).

[4] HHS also uses the data issuers submit to their EDGE servers for the calculation of the high-cost risk pool payments and charges added to the HHS risk adjustment methodology beginning with the 2018 benefit year.

for the purposes of risk adjustment transfer calculations for states where HHS operates the risk adjustment program. The purpose of HHS–RADV is to ensure issuers are providing accurate and complete risk adjustment data to HHS, which is crucial to the purpose and proper functioning of the HHS-operated 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 HHS-operated risk adjustment program.[5] 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[6] 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 errors discovered as a 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 (IVA) of a sample of risk adjustment data selected by HHS.[7] The issuer provides demographic, enrollment, and claims data and medical record documentation for a sample of enrollees selected by HHS to its IVA entity for data validation. After the IVA entity has validated the HHS-selected sample, a subsample is validated in a second validation audit (SVA).[8] The SVA is conducted by an entity HHS retains to verify the accuracy of the findings of the IVA.

HHS conducted two pilot years of HHS–RADV for the 2015 and 2016 benefit years[9] 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 pools. The 2017 benefit year HHS–RADV was the first payment year that resulted in adjustments to issuers' risk scores and the risk adjustment transfers in the applicable state market risk pools as a result of HHS–RADV findings.[10] [11]

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).[12] The 2013 RADV White Paper discussed and sought comment on a number of potential considerations for the development and operation of HHS–RADV. Based on the feedback received, HHS promulgated regulations to implement HHS–RADV that we have modified in certain respects based on experience and public input, 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.

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. In the March 11, 2014 **Federal Register** (79 FR 13743), we published the 2015 Payment Notice final rule, which finalized HHS–RADV requirements related to sampling; IVA standards, SVA 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. 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 IVA 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

[5] HHS also has general authority to audit issuers of risk adjustment covered plans pursuant to 45 CFR 153.620(c).

[6] See 45 CFR 153.20 for the definition of ''risk adjustment covered plan.''

[7] 45 CFR 153.630(b).

[8] 45 CFR 153.630(c).

[9] 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.*

[10] 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.*

[11] 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.

[12] 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.*

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 adjust issuers' risk scores when an issuer's failure rate is materially 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 IVA 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,[13] 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 the 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.[14]

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 **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) (2020 Payment Notice). The final rule included policies related to incorporating risk adjustment prescription drug categories (RXCs)[15] into HHS–RADV beginning with the 2018 benefit year and extending the Neyman allocation to the 10th stratum for HHS–RADV sampling. We also finalized using precision analysis to determine whether the SVA results of the full sample or the subsample (of up to 100 enrollees) results should be used in place of IVA results when an issuer's IVA 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 HHS 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 rule, 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 EDGE HCCs (hierarchical condition categories) within an HCC failure rate group would not be determined to be 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,[16] 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.[17] Although the statutory provisions for many other PPACA programs appropriated funding, authorized amounts to be appropriated, or provided budget authority in advance

[13] 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 through 17504.

[14] "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.*

[15] 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.

[16] See, *e.g.,* 78 FR 15441 and 83 FR 16930.

[17] Also see *New Mexico Health Connections* v. *United States Department of Health and Human Services,* 946 F.3d 1138 (10th Cir. 2019).

of appropriations,[18] 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.[19] Indeed, unlike the Medicare Prescription Drug, Improvement and Modernization Act of 2003, 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.[20] Congress did not give HHS discretion to implement a risk adjustment program 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. Prior to the proposed rule, 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).[21] We considered comments received on the 2019 RADV White Paper and in connection with previous rules as we developed the policies in the proposed rule. For this final rule, we considered all public input we received on the topics addressed in the proposed rule as we developed the finalized policies.

II. Provisions of the Final Regulations and Analyses and Responses to Public Comments

In the June 2, 2020 **Federal Register** (85 FR 33595), we published the ''Amendments to the HHS-Operated Risk Adjustment Data Validation Under the Patient Protection and Affordable Care Act's HHS-Operated Risk Adjustment Program'' proposed rule. The proposed rule proposed several refinements to the HHS–RADV error rate calculation, and proposed to transition away from the current prospective application of HHS–RADV results.[22] The proposals were designed to specifically address stakeholder feedback received after the first payment year of HHS–RADV. In addition to soliciting comments on the specific policy proposals in the proposed rule, we requested feedback on the potential impact of the COVID–19 public health emergency on the proposed effective dates for implementation of the proposals. We received 25 comments from health insurance issuers, industry trade associations, and other stakeholders. These comments ranged from general support of or opposition to the proposed changes to specific questions or comments regarding proposed changes. We also received a number of comments and suggestions that were outside the scope of the proposed rule that are not addressed in this final rule. In this final rule, we provide a summary of the proposed changes, a summary of the public comments received that directly relate to these proposals, our responses to these comments, and a description of the provisions we are finalizing.

This rule finalizes the proposed changes to two aspects of HHS–RADV: (A) The error rate calculation, and (B) the application of HHS–RADV results, with the modifications described below. Beginning with the 2019 benefit year of HHS–RADV,[23] we are finalizing as proposed the following refinements to the error rate calculation: (1) An adjustment to the HCC grouping methodology to address the influence of the HCC hierarchies and coefficient estimation groups; (2) a sliding scale adjustment for calculating an issuer's adjustment factor that changes the confidence intervals for determining outliers and applies a sliding scale adjustment in cases where an outlier issuer is close to the edges of the confidence interval for one or more HCC failure rate groups; and (3) a modification to the error rate calculation in cases where a negative error rate outlier issuer also has a negative failure rate. We are also finalizing the transition from the current prospective application of HHS–RADV results [24] to an approach that would apply HHS–RADV results to the benefit year being audited. After consideration of comments, we will switch to the concurrent application of HHS–RADV results beginning with the 2020 benefit year.[25] We believe these policies address stakeholder feedback received and our experience with the first payment year of HHS–RADV on these issues. These finalized policies seek to further the integrity of HHS–RADV while maintaining stability, promoting fairness and improving the predictability of HHS–RADV. The following is a summary of the comments received on the proposed rule's timeline for implementing these policies: [26]

Comments: One commenter was concerned that the COVID–19 public health emergency would impact the completeness of 2019 (and possibly 2020) data while another commenter

[18] 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).

[19] See 42 U.S.C. 18063.

[20] 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).

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

[22] The exception to the current prospective application of HHS–RADV results is for exiting issuers identified as positive error rate outliers, whose HHS–RADV results are applied to the risk scores and transfer amounts for the benefit year being audited. See the 2020 Payment Notice, 84 FR at 17503–17504.

[23] As part of the Administration's efforts to combat the Coronavirus Disease 2019 (COVID–19), we announced the postponement of the 2019 benefit year HHS–RADV process. See *https://www.cms.gov/files/document/2019-HHS-RADV-Postponement-Memo.pdf.* Also, we have provided further guidance on the updated schedule for the 2019 benefit year HHS–RADV, which is outlined in the 2019 Benefit Year Timeline of Activities: *https://www.regtap.info/uploads/library/HRADV_Timeline_091020_5CR_091020.pdf.*

[24] The exception to the current prospective application of HHS–RADV results is for exiting issuers identified as positive error rate outliers, whose HHS–RADV results are applied to the risk scores and transfer amounts for the benefit year being audited.

[25] As detailed in section II.B, to effectuate the transition beginning with the 2020 benefit year, we will aggregate results from the 2019 and 2020 benefit years of HHS–RADV for non-exiting issuers using the average error rate approach and apply the aggregated results to 2020 risk scores and transfers.

[26] We note that a correction notice was issued for the proposed rule to address the misalignment of certain text between the final draft version of the proposed rule approved for publication and the published version in the **Federal Register**. See 85 FR 38107 (June 25, 2020). Since publishing the correction notice, an additional error between the two versions was identified. When describing the current HHS–RADV error methodology in the proposed rule at 85 FR 33599, the upper bound of the confidence interval was incorrectly published as $UB_G = \mu\{GFR_G\} - sigma_cutoff * Sd\{GFR_G\}$. This formula should have instead been published as $UB_G = \mu\{GFR_G\} + sigma_cutoff * Sd\{GFR_G\}$.

expected COVID–19 to affect chart retrieval and provider documentation within the chart. One commenter did not see a need to further delay the stabilizing measures in the proposed rule due to COVID–19.

Response: Recognizing the need for providers and provider organizations to focus exclusively on caring for patients during the COVID–19 public health emergency, we postponed the start of 2019 benefit year HHS–RADV activities.[27] As recently announced, IVA samples for 2019 benefit year HHS–RADV will be released in January 2021 and we anticipate 2020 benefit year HHS–RADV will commence as usual with the release of IVA samples in May 2021.[28] We continue to monitor the COVID–19 pandemic, including potential medical record retrieval issues and will consider whether additional flexibilities for HHS–RADV are appropriate. However, we are not codifying or finalizing any specific COVID–19 policies in this rulemaking.

Comments: Some commenters who supported the proposed error rate calculation changes asked HHS to also apply the changes to the 2017 and 2018 benefit years of HHS–RADV. A different commenter opposed applying the proposed changes starting with the 2019 benefit year HHS–RADV, expressing the belief it would be retroactive to do so, and instead supporting the adoption of these proposals for future benefit years. Other commenters supported policies in the rule applying beginning with the 2019 benefit year.

Response: The policies being finalized in this rule only impact the calculation of error rates and the application of the HHS–RADV results that occur at the end of the HHS–RADV process. Because the 2019 benefit year of HHS–RADV has not begun [29] and, under the updated timeline, the calculation of the error rates for 2019 benefit year of HHS–RADV will not occur until February 2022, we disagree that applying the error rate calculation refinements finalized in this rule to the 2019 benefit year would be retroactive. Further, for the reasons outlined in the proposed rule and this rule, we believe these refinements are important and should be applied as soon as practicable. However, we believe that application of this rule to 2017 and 2018 benefit years of HHS–RADV would not be appropriate because the applicable error rate calculations are complete.[30] [31] We are therefore applying the error rate calculation modifications finalized in this rule beginning with the 2019 benefit year of HHS–RADV, as proposed. Similarly, for the application of HHS–RADV results, in light of the delay of 2019 benefit year HHS–RADV and for the reasons outlined below in Section II.B., we are finalizing the policy to begin applying HHS–RADV results to the benefit year audited beginning with the 2020 benefit year which is as soon as practicable.[32]

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,[33] 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 metric. HHS defines the national statistically meaningful metric as the weighted mean and standard deviation of the failure rate calculated based on all issuers' HHS–RADV results. Each issuer's failure rates are compared to these national metrics to determine whether the issuer's failure rate is an outlier. Based on outlier issuers' failure rate results, their error rates are calculated and applied to their plan liability risk scores.[34]

In response to comments received on the 2019 RADV White Paper and to help put the proposed changes in context, the proposed rule outlined the current error rate calculation methodology.[35] This included 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.

Consistent with 45 CFR 153.350(c), HHS applies the outlier issuer's error rate to adjust that issuer's applicable benefit year plan liability risk score.[36] This risk score change, 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 risk adjustment 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 proposed 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 subsequently adjusted transfers. We also proposed changes to mitigate the impact of HHS–RADV

[27] *https://www.cms.gov/files/document/2019-HHS-RADV-Postponement-Memo.pdf.*

[28] See the ''2019 Benefit Year HHS–RADV Activities Timeline'' *https://www.regtap.info/uploads/library/HRADV_Timeline_091020_5CR_091020.pdf.*

[29] As noted above, the start of the 2019 benefit year HHS–RADV process was postponed until the 2021 calendar year due to the COVID–19 public health emergency.

[30] See the 2017 HHS–RADV timeline, available at: *https://www.regtap.info/uploads/library/HRADV_JobAid_timeline_5CR_032819.pdf;* and *https://www.regtap.info/uploads/library/HRADV_Timeline_073119_5CR_120219.pdf.* Also see the 2018 HHS–RADV timeline, available at: *https://www.regtap.info/uploads/library/HRADV_Timeline_030420_V1_RETIRED_5CR_041320.pdf.*

[31] See the 2017 and 2018 HHS–RADV results memos, available at: *https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/2017-Benefit-Year-HHS-Risk-Adjustment-Data-Validation-Results.pdf* and *https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/2018_BY_RADV_Results_Memo.pdf.*

[32] As detailed below, to effectuate the transition beginning with the 2020 benefit year, we will aggregate results from the 2019 and 2020 benefit years of HHS–RADV for non-exiting issuers using the average error rate approach and apply the aggregated results to 2020 benefit year risk scores and transfers.

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

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

[35] See 85 FR at 33599–33600. Also see, *supra,* note 26.

[36] Exiting positive error rate 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. As detailed in Section II.B, we are finalizing the proposed transition from the prospective application of HHS–RADV results such that risk score error rates will also be applied to the benefit year being audited beginning with the 2020 benefit year of HHS–RADV for non-exiting issuers.

adjustments that result from negative error rate outlier issuers with negative failure rates. After consideration of comments, we are finalizing the refinements to the error rate calculation, as proposed, beginning with the 2019 benefit year of HHS–RADV. These targeted policies are intended as interim, incremental measures while we continue to analyze HHS–RADV results and consider potential further refinements and changes to the HHS–RADV methodology, including potential significant changes to the outlier determination process and the error rate 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.[37] For risk adjustment model development, this includes: (1) The hierarchies of HCCs, (2) HCC coefficient estimation groups, (3) *a priori* stability constraints, and (4) hierarchy violation constraints. For HHS–RADV, medical conditions are grouped for the HHS–RADV HCC failure rate groups. These grouping processes are not concurrent. More specifically, the grouping processes related to model development are implemented prior to the benefit year and the HHS–RADV HCC failure rate groups are implemented after the benefit year. Our experience in the initial years of HHS–RADV found that differences among these grouping processes interact in varying ways and may result in greater or lesser HHS–RADV adjustments than may be warranted in certain circumstances.

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 [38] 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.[39] Condition categories become HCCs once these hierarchies are imposed.

As noted previously, 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 the risk adjustment model and plan liability risk score calculation. Although HCCs reflect hierarchies among related disease categories, multiple HCCs can accumulate for enrollees with unrelated diseases; 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 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 such that the coefficients of two or 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,[40] 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.

Although some of these HCC coefficient estimation groups occur within hierarchies, some HCC coefficient estimation groups include HCCs that do not share a hierarchy. 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 child models. The adult model HCC coefficient estimation groups for the V05 risk adjustment models [41] 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

[37] See 85 FR at 33601.

[38] 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 at 29175.

[39] 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 June 17, 2019 document ''Potential Updates to HHS–HCCs for the HHS-operated Risk Adjustment Program'' (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#page=11.*

[40] As described in the ''Potential Updates to HHS–HCCs for the HHS-operated Risk Adjustment Program'' Paper, available at ''*https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Potential-Updates-to-HHS-HCCs-HHS-operated-Risk-Adjustment-Program.pdf#page=11.*

[41] The shorthand ''V05'' refers to the current HHS–HCC classification for the HHS risk adjustment models, which applies through the 2020 benefit year. V07 is the HHS–HCC classification for the HHS risk adjustment models, which applies beginning with the 2021 benefit year.

TABLE 1—HCC COEFFICIENT ESTIMATION GROUPS FROM ADULT RISK ADJUSTMENT MODELS V05—Continued

HHS HCC	V05 HHS–HCC label	Adult model HCC coefficient estimation group
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
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. These *a priori* stability 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. Prior to the 2021 benefit year recalibration, only one *a priori* stability constraint was applied to the models, and this constraint was only applied to the child models.[42]

HCC coefficient estimation groups 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 than a higher severity HCC in the same hierarchy). 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. To create the HCC failure rate groupings for HHS–RADV, 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 frequency 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.[43]

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 ways that could result in misalignments.[44]

Based on HHS's initial analysis of the 2017 benefit year HHS–RADV results, and in response to comments to the 2019 RADV White Paper, HHS considered an option in the proposed rule to address the influence of the HCC hierarchies and HCC coefficient estimation groups on the HCC failure rate groupings in HHS–RADV. We proposed to modify the creation of HHS–RADV HCC failure rate groupings to 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 proposed that, when HHS calculates EDGE and IVA frequencies for each individual HCC, we would aggregate HCCs that are in the same HCC coefficient estimation group

[42] In the 2021 Payment Notice (85 FR at 29178), we finalized an additional *a priori* stability constraint to the child 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, and revised the single transplant stability constraint in the child models to be two stability constraints to better distinguish transplant cost differences.

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

[44] See 85 FR at 33603–33604. Also see Section 3.3 of the 2019 RADV White Paper.

in the adult risk adjustment models (and, therefore, have coefficients constrained to be equal to one another) into one ''Super'' HCC, prior to calculating individual HCC failure rates and sorting the HCCs into low, medium, and high failure rate groups for HHS–RADV. These new frequencies, including the aggregated frequencies of HCC coefficient estimation groups and the individual frequencies of all other HCCs that are not aggregated with other HCCs because they are not in any coefficient estimation groups, would be considered frequencies of ''Super HCCs.''

Under the proposed methodology, we would modify the current HCC failure rate grouping methodology as follows:

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

Where:

c is the index of the cth Super HCC;

$freqEDGE_h$ 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;

$freqEDGE_c$ 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 $freqEDGE_h$;

$freqIVA_h$ 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;

$freqIVA_c$ 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 would be grouped according to the current HHS–RADV failure rate grouping methodology.

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 independently and additively to an issuer's failure rate in a HCC failure rate grouping. This proposal would refine the current methodology to better identify and focus HCC failure rates used in outlier determination on actual differences in risk and costs. Our tests of this proposed policy on HHS–RADV results data revealed that between an estimated 85.2 percent (2018 data) and 98.1 percent (2017 data) of the occurrences of HCCs on EDGE belong to HCCs that would be assigned to the same failure rate groups under the proposed ''Super HCC'' methodology as they have been under the current methodology as seen in Table 2. Although the impact on individual issuer results may vary depending upon the accuracy of their EDGE data submissions and the rate of occurrence of various HCCs in their enrollee population, the national metrics used for HHS–RADV, that is, the weighted means and weighted standard deviations, would only be slightly affected, as seen in Table 3. 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 would have similar failure rates under both methodologies. 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.

TABLE 2: Estimated Changes in the HCC Groupings Using Super HCCs Based on Aggregates of Adult Model HCC Coefficient Estimation Groups (Using Prior Benefit Year HHS-RADV Results Data)

Unique Count of HCCs in Each Failure Rate Group		New Methodology					
		2017 Data			2018 Data		
		Low	Medium	High	Low	Medium	High
Current Methodology	Low	31	1	1	27	5	1
	Medium	2	29	4	1	27	4
	High	1	5	53	1	7	55
Frequency of HCC Occurrences on EDGE		**New Methodology**					
		2017 Data			**2018 Data**		
		Low	Medium	High	Low	Medium	High
Current Methodology	Low	32.2%	0.0%	0.0%	26.9%	5.0%	0.0%
	Medium	0.1%	33.0%	1.0%	5.0%	27.0%	1.3%
	High	0.3%	0.3%	32.9%	0.4%	3.2%	31.3%

In testing this proposal to create Super HCCs in HHS–RADV, we grouped HCCs in the same HCC coefficient estimation group in the adult risk adjustment models. We chose to use the adult risk adjustment models for testing because the majority of the population with HCCs in the HHS–RADV samples are subject to the adult models (88.3 percent for the 2017 benefit year; 89.1 percent for the 2018 benefit year).[45] 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 accuracy of HCC failure rate estimates across the entire population in aggregate. Therefore, we proposed to use HCC coefficient estimation groups in the adult risk adjustment models to define Super HCCs for all HHS–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 did not consider including *a priori* stability constraints or hierarchy violation constraints in the aggregation of Super HCCs.[46] We also did not consider hierarchy violation constraints as a part of the sorting algorithm in order to balance complexity and consistency. For example, if, in a given benefit year, the magnitudes of two coefficients that share a hierarchy happen to decrease in order of their conditions' theoretical severity, the coefficients would violate the assumptions of the hierarchy structure and would be subject to a hierarchy violation constraint in that year's risk adjustment models. However, if the magnitude of those two coefficients increase in the order of their conditions' severity in the subsequent year, as would generally be expected, the coefficients would be consistent with the assumptions of the hierarchy structure and would not be constrained to be equal as a part of a hierarchy violation constraint. Because these year-to-year changes in hierarchy violation constraints are based solely on the magnitude of each year's initial coefficient estimates, using them in the grouping of Super HCCs would make those groupings less stable and transparent, and would reduce predictability for issuers.

Due to these considerations, we proposed to combine HCCs into Super HCCs defined only by HCC coefficient estimation groups in the adult risk adjustment models 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. As proposed, these Super HCC groupings would apply to all HHS–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 solicited comment on all aspects of this proposal. We also solicited comments on whether, in addition to the Super HCCs based on the adult risk adjustment models, HHS should create separate infant Super HCCs for each maturity and severity type in the infant risk adjustment models. Additionally, we solicited comments on whether we should consider incorporating *a priori* stability constraints from the child models or hierarchy violation constraints from the adult models when defining Super HCCs.

After consideration of the comments received, we are finalizing this policy as proposed, and will combine HCCs in HCC coefficient estimation groups in the adult risk adjustment models, which effectively have equal coefficients, into Super HCCs prior to sorting the HCCs into low, medium and high failure rate groups for HHS–RADV. This refinement to the error rate calculation will apply starting with the 2019 benefit year of HHS–RADV. These Super HCC groupings will apply to all HHS–RADV sample enrollees, regardless of the risk adjustment models to which they are subject. Therefore, although the aggregation will be based upon the adult models, enrollees subject to the child and infant models will have their HCCs included in the aggregated counts when they have an HCC that is listed as sharing a coefficient estimation group with other HCCs in the adult models. The resulting Super HCCs will then be sorted into high, medium, and low failure rate groups using the sorting process described in the applicable benefit year's HHS–RADV Protocols.[47] 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, will be grouped according to the current HHS–RADV failure rate grouping methodology.

Comments: All comments on this policy supported the proposal to adjust the HCC failure rate grouping methodology to define Super HCCs based upon the HCC coefficient estimation groups in the adult risk adjustment models. Several commenters requested we expand the proposed definition of Super HCCs to include the grouping of conditions used to create the variables for the infant models. Some of these commenters added that implementing this expansion for the infant models should be done in a way that avoids year-to-year stability concerns, if possible, while other comments requested that we publish an analysis on the impacts of such an expansion prior to implementing it.

In addition, some commenters agreed that the inclusion of *a priori* stability constraints from the child models would be inappropriate due to their additive nature, with a few of these commenters also agreeing that hierarchy violation constraints should not factor into the definitions of Super HCCs. However, other commenters requested that HHS include HCCs involved in a hierarchy violation constraint in the same Super HCC. Some commenters requested we publish an analysis on including *a priori* stability constraints as part of the process to create Super HCCs.

Response: We are finalizing the refinement to the HCC failure rate grouping methodology as proposed and will place all HCCs that share an HCC coefficient estimation group in the adult risk adjustment models into the same HCC failure rate grouping beginning with the 2019 benefit year of HHS–RADV. Although the aggregation will be based upon the adult models, the child

[45] For 2017, this was calculated after removing issuers in Massachusetts and incorporating cases where issuers failed pairwise and the SVA sub-sample was used.

[46] Both *a priori* stability constraints and hierarchy violation constraints are described earlier in this section (Section II.A.1) of the rule. Also see 85 FR at 33602–33603.

[47] See Section 11.3.1 of the 2018 HHS–RADV Protocols at *https://www.regtap.info/uploads/library/HRADV_2018Protocols_070319_RETIRED_5CR_070519.pdf* for a description of the process prior to the introduction of Super HCCs. Beginning with the 2019 benefit year of HHS–RADV, Super HCCs would take the place of HCCs in the process. The 2019 HHS–RADV Protocols have thus far only been published in part at *https://www.regtap.info/uploads/library/HRADV_2019_Protocols_111120_5CR_111120.pdf.* The section of the 2019 HHS–RADV Protocols pertaining to HCC grouping for failure rate calculations is not included in the current version. Once published, this section will be updated to include steps related to creation of Super HCCs.

and infant models will have their HCCs included in the aggregated counts when they have an HCC that is listed as sharing a coefficient estimation group with other HCCs in the adult models. As explained in the proposed rule and in this rule, we believe this change mitigates the misalignments that occur when HCCs with the same risk score coefficient are sorted into different HCC failure rate groupings while increasing the stability of year-to-year HCC failure rate grouping assignments. To promote fairness and ensure the integrity of the program, we do not believe that a RADV finding that reflects an EDGE data miscoding of one condition as another condition from the same coefficient estimation group should contribute to any of an issuer's three failure rates. This refinement to the HHS–RADV failure rate grouping methodology ensures that these types of HCC miscodings with no risk score impact do not impact an issuer's HHS–RADV error rate.

We appreciate the comments about the creation of separate infant Super HCCs and investigated the potential adoption of separate infant model terms. Our analysis found that such an approach would likely result in more year-to-year uncertainty and instability due to the relatively small sample size for some infant model terms—notably, only around 5 percent of 2017 [48] and 2018 HHS–RADV sample enrollees in strata 1 through 9 with EDGE HCCs were infants. As a result, HCC counts and failure rates for potential infant-only Super HCCs would be more likely to vary due to random selection, yielding less year-to-year stability among HCC failure rate group assignments. Therefore, in the interest of stability, we believe that basing the definitions of Super HCCs on coefficient estimation groups from the adult risk adjustment models is more appropriate. As noted earlier, the majority of the population with HCCs in the HHS–RADV samples are subject to the adult models (88.3 percent for the 2017 benefit year; 89.1 percent for the 2018 benefit year).[49]

We also appreciate the comments regarding inclusion of hierarchy violation constraints when creating Super HCCs, such that HCCs involved in a hierarchy violation constraint would be included in the same Super HCC. As explained in the proposed rule, we did not consider hierarchy violation constraints when developing the Super HCC proposal in order to balance complexity and consistency, since these constraints can change from year-to-year as a natural result of the annual recalibration updates to the model coefficients. Similar to the concerns for the separate infant model Super HCCs, these year-to-year changes would make HCC groupings for these HCCs less stable and transparent, and would reduce predictability for issuers. Further, we note that hierarchy violation constraints may occur in a single metal-level and age group in just one of the three data years used to create the blended coefficients. For example, the 2021 benefit year coefficients reflect a weighted average of coefficients calculated separately from 2016, 2017, and 2018 benefit year EDGE data. If there is a hierarchy violation among three HCCs that share a hierarchy in the silver adult model fitted to 2018 EDGE data, a hierarchy violation constraint would be applied to the three coefficients calculated from that data set alone, excluding any coefficients from the 2016 and 2017 benefit years, and any other metal levels and age groups from the 2018 benefit year. As a result, when the coefficients from the separate data years are blended, the hierarchy violation constraint may not be apparent in the final coefficients and the final coefficients for the HCCs in the affected hierarchy may differ from one another.

Additionally, even if a hierarchy violation constraint is necessary for the same hierarchy in all three data years, and is therefore apparent in the final risk adjustment coefficients, the hierarchy violation constraint could involve a very small number of enrollees specific to a particular metal level and age group model (for example, the gold metal level child model). Although the coefficients involved in such a hierarchy violation constraint would all be equal to one another, the coefficients from age group models unaffected by hierarchy violation constraints are likely to differ according to the severity of the HCCs in the hierarchy, and it would be appropriate to capture the resulting risk score differences in HHS–RADV. Therefore, a methodology that included hierarchy violation constraints in the definition of Super HCCs would have to keep the relevant HCCs in the applicable metal level and age group model affected by the hierarchy violation constraints separate from the same HCCs in metal levels and age group models that are unaffected. This would result in individual Super HCCs dedicated to only the HCCs affected by a given hierarchy violation constraint from HHS–RADV sample enrollees subject to the affected metal level and age group model. As such, the individual Super HCC failure rate calculation for that hierarchy violation constraint would be based on a very small sample, leading to instability for the HCC failure rate group assignment for that hierarchy violation constraint. It would also increase the complexity associated with adoption of this refinement to the HCC failure rate grouping methodology. In contrast, coefficient estimation groups are consistent across all five metal level adult models, and are almost identical to the coefficient estimation groups across all five metal level child models. As such, it is much more appropriate to define Super HCCs for all enrollees based on the adult coefficient estimation groups, because nearly all enrollees with an EDGE miscoding between two HCCs in a coefficient estimation group would be assigned the same risk score for either HCC. This consistency allows us to utilize a much larger sample size during the calculation of Super HCC-specific failure rates, namely, the entire HHS–RADV sample, resulting in more stable failure rate estimates and HCC failure rate group assignments. Defining Super HCCs based on the adult coefficient estimation groups is also easy to implement as an interim measure to address the identified misalignment that occurs in situations where HCCs in the same HCC coefficient estimation group are sorted into different HCC failure rate groupings.

Finally, we appreciate the comments requesting more analysis on including *a priori* stability constraints from the child models in the definition of Super HCCs. For similar reasons to those noted in the discussion of the hierarchy violation constraints and variables from infant models, including *a priori* stability constraints from the child models in the definition of Super HCCs would result in very small sample sizes for the purposes of determining the Super HCC-level failure rate prior to sorting into HCC failure rate groups. As such, our analysis of the inclusion of *a priori* stability constraints for the child models found that it would likely result in less year-to-year uncertainty in that model than basing Super HCCs on coefficient estimation groups alone. Moreover, HCCs subject to *a priori* stability constraints are additive in the risk adjustment models, whereas HCCs within coefficient estimation groups are not.[50] This difference is due to the fact

[48] For 2017, this was calculated after removing issuers in Massachusetts and incorporating cases where issuers failed pairwise agreement and the SVA sub-sample was used.

[49] Ibid.

[50] The additive nature of HCCs subject to *a priori* stability constraints as opposed to other groupings

that many of the *a priori* stability constraints reflect unrelated conditions, and therefore, a miscoding of one HCC within an *a priori* stability constraint would not be expected to impact the likelihood that another HCC in that *a priori* stability constraint would also be miscoded. In contrast, coefficient estimation groups reflect related conditions that could conceivably be miscoded as one another on EDGE. Therefore, we do not believe that it is appropriate to include *a priori* stability constraints from the child models in the definition of Super HCCs.

Comments: A few commenters supported the proposed changes as valuable interim measures, but stated that the HCC failure rate grouping methodology may require additional improvements in the future and asked that HHS continue to analyze and propose refinements to the HCC grouping process for HHS–RADV. Some of these commenters emphasized that stability of HCC failure rate group assignment from year-to-year should be a priority when considering potential future changes.

Response: We appreciate these comments. As noted in the proposed rule, the Super HCC refinement is intended to address the misalignment that occurs in situations where HCCs in the same HCC coefficient estimation group are sorted into different HCC failure rate groupings on an interim basis while we continue to assess different longer-term options. We remain committed to ensuring the integrity and reliability of HHS–RADV and agree that year-to-year stability is an important factor to consider when analyzing potential future changes. We continue to explore potential modifications to this program, including to the HCC grouping methodology, for future benefit years and will propose any such changes through notice-and-comment rulemaking.

Comments: Several commenters requested that HHS release more information about the HCC failure rate grouping proposal to create Super HCCs. This included requests for more information about the degree to which validation failures relate to hierarchies for 2018 HHS–RADV, analysis on year-to-year stability, and a further explanation of the proposed refinement to the HCC failure rate grouping methodology.

Response: Once the data became available, we conducted an additional analysis of the Super HCC proposal using 2018 benefit year HHS–RADV results. This further analysis provided roughly the same figure for the proportion of newly identified HCCs which could be attributed to a miscoding of an HCC in the same hierarchy, or in the same coefficient estimation group, as the analysis of 2017 benefit year HHS–RADV results used to develop the Super HCC proposal, namely, about 1/3rd of newly identified HCCs. Among non-validated HCCs, the rate that could be attributed to miscoding of an HCC in the same hierarchy was slightly higher in our analysis of 2018 data (about 1/7th of non-validated HCCs) than it was for 2017 data (about 1/8th of non-validated HCCs). Additionally, in response to comments, we note that in both 2017 and 2018 HHS–RADV results, approximately 1/3rd of HCCs that could be attributed to miscoding of an HCC in the same hierarchy also shared a coefficient estimation group.[51] The refinement to the HCC failure group rate methodology finalized in this rule will ensure that these HCCs will have no impact on failure rates. More specifically, adoption of this change for HCCs in the same coefficient group ensures they are not sorted into different HCC failure rate groupings and avoids making HHS–RADV adjustments to risk scores when they are not conceptually warranted.

In response to the comments, we also provide the following additional example regarding the calculation of a Super's HCC failure rate using $freqEDGE_c$, $freqIVA_c$, and FR_c values for Super HCCs.[52] HCC 54 *Necrotizing Fasciitis* and HCC 55 *Bone/Joint/Muscle Infections/Necrosis* share a HCC coefficient estimation group, and therefore those HCC failure rates would be grouped together to form a Super HCC. For example, if $freqEDGE_{h54}$ is 30 and $freqEDGE_{h55}$ is 70, nationally, and if $freqIVA_{h54}$ is 15 and $freqIVA_{h55}$ is 65, nationally, then $freqEDGE_{c54\&55}$ is 100 and $freqIVA_{c54\&55}$ is 80, yielding $FR_{c54\&55} = 1 - 80/100 = 20\%$. This is in contrast to cases such as HCC 1 *HIV/AIDS,* which does not share a coefficient estimation group with any other HCCs. In this second example, $freqEDGE_c$ will be equal to $freqEDGE_h$, $freqIVA_c$ will be equal to $freqIVA_h$, and FR_c will be equal to FR_h, the value of the national failure rate for HCC 1.

As explained in the proposed rule, after the calculation of $freqEDGE_c$, $freqIVA_c$, and FR_c, we will sort the 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—using the sorting process under the current HHS–RADV failure rate grouping methodology. The sorting process and failure rate grouping methodology are described in the HHS–RADV Protocols.[53] Specifically, HHS will calculate the HCC failure rate group for each Super HCC using the following method:

- Create a list containing each Super HCC and its associated failure rate.
- Sort Super HCCs from lowest to highest failure rate (FR_c).
- Put the Super HCC with the lowest failure rate in the low failure rate group, and update the size of this group ($freqEDGE_{low}$) so that it is equal to $freqEDGE_{c1}$, that is, the value of $freqEDGE_c$ for the first Super HCC from the sorted list. Put the next Super HCC from the sorted list in the low failure rate group, and update the group size to $freqEDGE_{low} + freqEDGE_{ci}$, the value of $freqEDGE_c$ for the *i*-th Super HCC from the sorted list. Repeat this sorting process until the size of $freqEDGE_{low}$ reaches or exceeds 1/3rd of the total frequency of HCCs recorded on EDGE ($\Sigma freqEDGE_h$ across all HCCs, which is equal to $\Sigma freqEDGE_c$ across all Super HCCs).
- After the low failure rate group has reached the 1/3rd cut off, HHS will put the next Super HCC from the sorted list into the medium failure rate group, and will update the size of this group ($freqEDGE_{medium}$) so that it is equal to $freqEDGE_{ci}$. We will then put the next Super HCC from the sorted list into the medium failure rate group, and update the group size to $freqEDGE_{medium}$ +

of HCCs in the risk adjustment models is discussed in greater detail in the proposed rule (85 FR 33605). We have also previously discussed this feature of *a priori* stability constraints in the 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#page=11.*

[51] See Table 2 for a further comparison and analysis of the estimated changes reflecting implementation of the Super HCC refinement using 2017 and 2018 HHS–RADV data. Also see Tables 3 and 4 for a further analysis and comparison of the estimated changes reflecting implementation of the policies finalized in this rule using both 2017 and 2018 benefit year HHS–RADV results.

[52] Commenters should also refer to the illustrative example in the proposed rule. See 85 FR at 33605.

[53] See Section 11.3.1 of the 2018 HHS–RADV Protocols at *https://www.regtap.info/uploads/library/HRADV_2018Protocols_070319_RETIRED_5CR_070519.pdf* for a description of the process prior to the introduction of Super HCCs. Beginning with the 2019 benefit year of HHS–RADV, Super HCCs would take the place of HCCs in the process. The 2019 HHS–RADV Protocols have thus far only been published in part at *https://www.regtap.info/uploads/library/HRADV_2019_Protocols_111120_5CR_111120.pdf.* The section of the 2019 HHS–RADV Protocols pertaining to HCC grouping for failure rate calculations is not included in the current version. Once published, this section will be updated to include steps related to creation of Super HCCs.

$freqEDGE_{ci}$. We will repeat this process until $freqEDGE_{low} + freqEDGE_{medium}$ reaches or exceeds 2/3rds of the total number of HCCs recorded on EDGE ($\Sigma freqEDGE_{h}$ across all HCCs, which is equal to $\Sigma freqEDGE_{c}$ across all Super HCCs).

- The remaining Super HCCs, those with the highest failure rates, will then be assigned to the high failure rate group.

Because the inclusion of the final $freqEDGE_{ci}$ in a given failure rate group may result in the total frequency for that group going beyond 1/3rd of the total $\Sigma freqEDGE_{c}$, consistent with the current sorting process and methodology, HHS will then reexamine the HCC allocations between failure rate groups to ensure an even distribution of HCCs between failure rate groups such that each HCC failure rate group contains as close as possible to 1/3rd of the HCCs reported in EDGE. To accomplish this, we will first identify the final Super HCCs in the low and medium failure rate groups that result in a total $freqEDGE_{low}$ or $freqEDGE_{medium}$ that exceeds 1/3rd of the total $\Sigma freqEDGE_{c}$. Then we will generate multiple grouping scenarios such that the identified Super HCCs that cause $freqEDGE_{low}$ or $freqEDGE_{medium}$ to exceed 1/3rd of the total $\Sigma freqEDGE_{c}$ are instead included in the next higher failure rate group. These multiple grouping scenarios will contain all possible assignments of the two Super HCCs that cross the 1/3rd boundary for the low and medium failure rate groupings. For each grouping scenario, we will then calculate the potential values of $freqEDGE_{low}$, $freqEDGE_{medium}$, and $freqEDGE_{high}$ and then calculate the absolute distance between in each HCC failure rate group and 1/3rd. HHS will then choose the scenario that is closest to an exact 1/3rd split of HCC frequencies across groups. This scenario will be used as the final HCC failure rate grouping assignment for that HHS–RADV benefit year.

2. ''Payment Cliff'' Effect

The HHS–RADV error rate calculation methodology is based on the identification of outliers, as determined using certain national thresholds. Those thresholds are used to determine whether an issuer is an outlier and the error rate that will be used to adjust outlier issuers' risk scores. Under the current methodology, 1.96 standard deviations on both sides of the confidence interval around the weighted HCC group means are the thresholds 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.[54]

Some stakeholders have expressed concern that issuers with failure rates that are just outside of the confidence intervals receive an adjustment to their risk scores, even though these issuers' failure rates may not be significantly different from the failure rates of issuers just inside the confidence intervals who receive no risk score adjustment, creating a ''payment cliff'' or ''leap frog'' effect. For example, an issuer 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. 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 to identify outliers, some stakeholders suggested that HHS could mitigate 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.

While HHS considered several possible methods to address the payment cliff,[55] we proposed to address the payment cliff by adding a sliding scale adjustment to the current error rate calculation, such that the adjustments applied would vary based on the outlier issuer's 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 help reduce the disparity of risk score adjustments by using a linear adjustment.[56] We proposed to make this modification beginning with 2019 benefit year HHS–RADV.

To apply the sliding scale adjustment, we proposed to modify the calculation of the group adjustment factor (GAF) by providing a linear sliding scale adjustment for issuers whose failure rates are near the point at which the payment cliff occurs. To implement this policy, we needed to select the thresholds of the range ($innerZ_{r}$ and $outerZ_{r}$) to calculate and apply the sliding scale adjustment.[57] In the proposed rule, we proposed to calculate and apply a sliding scale adjustment between the 90 and 99.7 percent confidence interval bounds (from +/− 1.645 to 3 standard deviations). Under this proposal, the determination of outliers in HHS–RADV for each HCC grouping would no longer be based on a 95 percent confidence interval or 1.96 standard deviations from the mean, and would instead be based on a 90 percent confidence interval or 1.645 standard deviations from the mean. Specifically, this approach would adjust the upper and lower bounds of the confidence interval to be at 1.645 standard deviations from the mean, meaning that issuers with group failure rates outside of the 90 percent confidence interval in any HCC failure rate group will have their risk scores adjusted. This would result in more issuers being considered outliers under this methodology than under the current methodology, which uses a 95 percent confidence interval to detect outlier issuers, but these additional outlier issuers would face smaller GAFs due to the application of the sliding scale.

To calculate the sliding scale adjustment, we proposed to add an additional step to the calculation of issuers' GAFs that takes into consideration the distance of their group failure rates (GFRs) to the confidence interval. The present formula for an issuer's GAF, $GAF_{G,i} = GFR_{G,i} - \mu\{GFR_{G}\}$ would be 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:

[54] 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.

[55] See, *e.g.,* section 4.4.4 and 4.4.5 of the 2019 RADV White Paper.

[56] In the 2020 Payment Notice, 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.

[57] In the 2019 RADV White Paper, we considered four different options for calculating and applying 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.

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

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: where $disZ_{G,i,r} = a * z_{G,i} + b_r$, 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 (from +/−1.645 to 3 standard deviations). This $disZ_{G,i,r}$ value will 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 between +/−1.645 to 3 standard deviations. 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 is 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 that is being finalized later in this final rule, the final formula for the group adjustment when an outlier issuer is subject to the sliding scale ($GAF_{G,i,r}$ above) would be simplified to:

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

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 HCC's risk score adjustment factor for an issuer that had a failure rate with a z score within the range of values (from +/−1.645 to 3 standard deviations) selected for the sliding scale adjustment (*innerZr* and *outerZr*). All other enrollee adjustment factors would be calculated using the current formula for the $GAF_{G,i,r}$. Under this approach, the above formulas would be implemented as follows:

If $freqEDGE_{G,i} \geq 30$, 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 $-3 \leq z_{G,i} < -1.645$ or $3 \geq z_{G,i} > 1.645$

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

If $freqEDGE_{G,i} < 30$ 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 $outerZ_r$ and 1.645 (or −1.645, for negative outliers) as the value of $innerZ_r$.

We sought comment on this proposal, including the proposed calculation of the sliding scale adjustment and the thresholds used to calculate and apply it. We also considered retaining the 95 percent confidence interval (1.96 standard deviations) as an alternative way to smooth the payment cliff. However, as noted in the proposed rule, while we recognize this option would also mitigate the payment cliff, we were concerned it would weaken the HHS–RADV program by reducing its overall impact and the magnitude of HHS–RADV adjustments to risk scores of outlier issuers.[58]

After consideration of comments received, we are finalizing the proposed sliding scale adjustment to smooth the payment cliff effect for those issuers whose failure rates are near the point at which the payment cliff occurs. We will calculate and apply a sliding scale adjustment between the 90 and 99.7 percent confidence interval bounds (from +/−1.645 to 3 standard deviations) beginning with 2019 benefit year HHS–RADV. For outlier issuers with failure rates more than 3 standard deviations from the mean, the GAF will not be impacted by the sliding scale adjustment, but will instead continue to be calculated as the difference between the weighted mean group failure rate and the issuers' group failure rate.

Comments: Some commenters supported the proposal to apply the sliding scale adjustment between the 90–99.7 percent confidence interval. Several commenters supported the adoption of a sliding scale adjustment but wanted to retain the current confidence intervals and start the adjustment at the 95 percent confidence interval. These commenters were concerned with the increased number of outliers under the proposed sliding scale adjustment, which would result in more risk adjustment transfers being impacted by HHS–RADV results, arguing this would reduce predictability and stability of HHS–RADV. Other commenters expressed concern about the identification of more outliers under the proposed sliding scale adjustment, arguing it would be more disruptive especially during COVID–19. Some commenters stated that they did not believe that identifying outliers at the proposed 90 percent confidence interval would more accurately capture issuers' actuarial risk and some thought the proposed 90 percent confidence interval could lead to an increase in ''false positives'' when identifying outliers. These commenters stated that the 95 percent confidence interval imposes a more robust confidence interval for identifying ''true outliers.''

Some commenters wanted HHS to calculate error rates based on the difference between the edge of the confidence intervals and the outlier issuer's failure rate (instead of the difference between the weighted group mean or a sliding scale adjustment and the outlier issuer's failure rate). However, these commenters also supported the adoption of a sliding scale adjustment starting at the 95 percent confidence intervals, if HHS were to finalize a sliding scale adjustment. One commenter wanted HHS to identify outliers and calculate their GAF based on state specific group means to address potential over and under adjustments of outlier issuers relative to their state-based competitors. One commenter supported the current methodology without a sliding scale adjustment, noting that the payment cliff effect resulted from the policy of only adjusting for outliers and that any measures to address the payment cliff would dampen the impact of HHS–RADV. Other commenters stated that it is appropriate for issuers who fall outside of the 99.7 percent confidence interval (beyond 3 standard deviations) to be assessed a full penalty. Another commenter, that supported the adoption of a sliding scale adjustment, expressed concerns that even with the proposed adjustment there would still be a payment cliff effect for issuers with very similar error rates. This commenter also asked HHS to address this effect for the current benefit year and beyond, as well as prior years, of HHS–RADV.

Response: We are finalizing the sliding scale approach for calculating an outlier issuer's error rate using modified group adjustment factors for issuers' group failure rates between 1.645 to 3 standard deviations from the mean on both sides of the confidence interval as proposed. We will apply this adjustment to the error rate calculation beginning with the 2019 benefit year of HHS–RADV. We believe that using a linear sliding scale adjustment will provide a smoothing effect in the current error rate calculation for issuers with failure rates just outside of the confidence interval of an HCC group and will retain the current significant adjustment to the HCC group weighted mean for issuers beyond three standard deviations. This approach ensures 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.

We appreciate the comments supporting an alternative sliding scale

[58] See 85 FR at 33608.

adjustment that would begin at 1.96 standard deviations. As detailed in the proposed rule, we recognize this alternative adjustment would also address the payment cliff and would provide stability by maintaining the current thresholds used in the error rate calculation. However, these benefits are outweighed by the concerns that such an adjustment would weaken HHS–RADV by reducing its overall impact and the magnitude of HHS–RADV adjustments to outlier issuer's risk scores. As noted previously, the sliding scale adjustment that is finalized in this rule will mitigate the payment cliff effect while not impacting the error rate calculation for those outlier issuers who are not close the confidence intervals.

While we did not propose adjusting issuers' error rates to the state-specific means, we considered such an approach in response to comments. However, we do not believe that using state-specific means would address the payment cliff in the current error rate methodology. We also have concerns about using national metrics to determine outliers and then switching to state-specific means to calculate the GAFs. In addition, the adoption of a state-specific approach to calculate the GAF could create other issues, if states have small sample sizes (that is, a small number of issuers participated in HHS–RADV), this would create less confidence in the state mean metric being used to adjust issuers, and would introduce new complexities as each state would have a different calculation for the GAF. We therefore decline to adopt such an approach in this final rule. We also considered adjusting to the confidence intervals,[59] but we have concerns that this option minimizes the impact of HHS–RADV adjustments on risk scores and risk adjustment transfers—including those outlier issuers with high error rates who are furthest away from the confidence intervals.

While any outlier threshold by definition has the risk of flagging false positives, and that risk may be slightly greater at the 90 percent confidence interval, we believe that the 90 percent confidence interval will better encourage issuers to ensure accurate EDGE data reporting and the risk of flagging false positives is mitigated by the fact that the adjustments to these issuers will be small since they will be subject to the sliding scale adjustment. Furthermore, while we understand the concerns that use of the 90 percent confidence interval will increase the number of outliers, we have found that the overall impact of the proposed approach on risk adjustment transfers is less than the current methodology despite the increased number of outliers. As discussed in the 2019 RADV White Paper, we tested various potential sliding scale adjustments between the 90 and 99.7 percent confidence interval bounds using 2017 HHS–RADV results.[60] We found that even though including issuers whose failure rates fell between 1.645 and 1.96 standard deviations from the mean would increase the number of outliers, the sliding scale adjustment lowers the overall impact of HHS–RADV adjustments to transfers and results in the distribution of issuers' error rates moving closer to zero compared to the current methodology.[61] We also tested this policy on the 2018 benefit year HHS–RADV data once it became available and found similar results. We found that the sliding scale adjustment option between 1.645 and 1.96 standard deviations generally resulted in lower overall impact of HHS–RADV adjustment to risk adjustment transfers and the distribution of issuers' error rates moving closer to zero compared to the current methodology. Furthermore, we believe that the 90 percent confidence interval will maintain the program integrity impact of HHS–RADV despite the estimated reduced impact of HHS–RADV on risk adjustment transfers using the 90 percent confidence interval, and we are not concerned that increasing the number of outliers will be more disruptive during the COVID–19 public health emergency. More importantly, we believe that using the 90 percent confidence interval will preserve 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 magnitude 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 methodology for calculating and applying the sliding scale adjustment provides a balanced approach to mitigating the payment cliff effect in the current methodology and disagree that adoption of the adjustment would reduce predictability and stability of HHS–RADV.

We recognize the sliding scale adjustment finalized in this rule does not eliminate the payment cliff because the identification of outliers will still be based on the establishment and use of thresholds. As noted earlier, we are finalizing the targeted policies in this rule, such as the sliding scale adjustment, as incremental refinements to the current error rate methodology to address stakeholder feedback and our experience from the first payment year of HHS–RADV on these issues. We will continue to consider other potential changes to the error rate methodology for future benefit years, including potential significant changes to the outlier determination process, and as part of that process, we will also consider whether additional measures are necessary or appropriate to further mitigate the impact of the payment cliff after we have experience with the sliding scale adjustment finalized in this rule.

We will apply the sliding scale adjustment beginning with the 2019 benefit year of HHS–RADV, as proposed. We believe that application of this rule to the 2017 and 2018 HHS–RADV would not be appropriate because the error rate calculations for those benefit years are complete.[62] Further, it would disrupt issuers' well-settled expectations with respect to the calculation of HHS–RADV error rates and adjustments if we were to extend this new policy to the 2017 and 2018 benefit years. In addition, there is no need to apply the sliding scale adjustment to the earlier benefit years because HHS–RADV was not conducted for the 2014 benefit year and HHS–RADV was treated as a pilot for the 2015 and 2016 benefit years.[63]

Comments: A few commenters noted that the increase in the number of issuers identified as outliers due to the introduction of the sliding scale adjustment could increase volatility by increasing the likelihood that an issuer would be an outlier in three HCC failure rate groups, leading to larger overall error rates despite the smaller GAF in each group, or by creating several negative outliers in one state market risk pool. One commenter, who was concerned about the increased number of outliers, noted that issuers can have a larger HHS–RADV adjustment under the proposed sliding scale adjustment than under the current methodology.

[59] See section 4.4.2 of the 2019 RADV White Paper.

[60] See section 4.4.5 and Appendix C of the 2019 RADV White Paper.

[61] Ibid.

[62] See, *supra,* notes 30 and 31.

[63] See FAQ ID 11290a (March 7, 2016) available at: *https://www.regtap.info/faq_viewu.php?id=11290* and HHS-Operated Risk Adjustment Data Validation (HHS–RADV)—2016 Benefit Year Implementation and Enforcement (May 3, 2017) available at: *https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/HHS-Operated-Risk-Adjustment-Data-Validation-HHS-RADV-%E2%80%93-2016-Benefit-Year-Implementation-and-Enforcement.pdf.*

Some commenters were concerned that this volatility from the increased number and type of outliers could increase premiums or adversely affect issuers' finanical planning.

Response: We recognize that the sliding scale adjustment finalized in this rule will result in more issuers being identified as outliers than the current methodology.[64] However, when testing various potential sliding scale adjustment options, we found that even though including issuers whose failure rates fell between 1.645 and 1.96 standard deviations from the mean would increase the number of outliers, the sliding scale adjustment we are finalizing in this rule lowers the overall impact of HHS–RADV adjustments to risk adjustment transfers and results in the distribution of issuers' error rates moving closer to zero compared to the current methodology.[65] Therefore, we do not believe that using the sliding scale adjustment starting with the 1.645 confidence interval will increase volatility or impact premiums more than the previous methodology. Instead, we believe that the sliding scale adjustment finalized in this rule will preserve 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 the incorporation of the sliding scale adjustment as proposed provides a balanced approach to mitigating the payment cliff effect.

Under the new confidence intervals with the sliding scale adjustment beginning at 90 percent finalized in this rule, it is possible for an issuer to fail more HCC groups resulting in larger error rates than the previous methodology or for there to be more negative error rate outliers in a state market risk pool compared to the current methodology. In those cases, outlier issuers could have a higher error rate, or non-outlier issuers could be impacted by more outliers in their state market risk pool than under the current methodology that does not include a sliding scale adjustment. However, failure rates for the issuers newly identified as outliers due to the adoption of the sliding scale adjustment would be between 1.645 to 1.96 standard deviations. Since these issuers' failure rates are closer to the mean, the increase in error rates based on outlier status in several HCC failure rate groups would likely be small and could potentially be offset by reduced transfers from other issuers with failure rates between 1.96 and 3 standard deviations in the same state market risk pool.

Comments: Some commenters expressed concern that issues other than actual HCC validation errors that impact the measurement of actuarial risk, such as medical record retrieval issues or incorrect provider coding, may contribute to the variance in failure rates, and that it is therefore not appropriate to adjust outlier issuers to the mean. Other commenters noted that changing the confidence intervals does not ensure that validation of HCCs that contribute to actuarial risk is accurately measured through HHS–RADV; these commenters supported maintaining the current confidence intervals.

Response: HHS–RADV validates risk based upon the enrollee's medical record which generally aligns with how the Medicare Advantage risk adjustment data validation (MA–RADV) program operates. Specifically, § 153.630(b)(7)(ii) requires that the validation of enrollee health status (that is, the medical diagnoses) occur through medical record review, that the validation of medical records include a check that the records originate from the provider of the medical services, that they align with the dates of service for the medical diagnosis, and that they reflect permitted providers and services. When an issuer fails to submit a medical record or has submitted an inaccurate medical record, the issuer has failed to validate the issuer's risk under our regulations. We do not treat these medical record issues differently than other errors that can occur in HHS–RADV nor would we treat them differently for purposes of calculating GAF using the weighted group mean.

While we are amending the calculation of the GAF, we did not propose and are not finalizing any changes to no longer use the mean in the calculation of the GAF. The purpose of the sliding scale adjustment is to mitigiate the payment cliff effect that was occuring by adjusting outlier issuers just outside the confidence interval to the weighted group mean. To ensure that the validation of HCCs that contribute to actuarial risk is accurately measured through HHS–RADV, we proposed the HCC failure rate grouping policy being finalized in this rule. That policy is another targeted refinement to the current methodology and it is focused on ensuring that miscoding of HCCs in the same coefficient estimation group with the same risk scores does not contribute to an issuer's group failure rate. Additionally, in this rule, we are finalizing the application of HHS–RADV results to the benefit year being audited in response to stakeholder concerns about changes in population and risk score between benefit years.

Comments: A commenter requested that HHS release prior HHS–RADV results and data if the sliding scale adjustment policy is finalized.

Response: Summary information on issuers' 2017 and 2018 benefit years HHS–RADV results are available on the Premium Stabilization Program page of the CCIIO website, which can be accessed at *https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs.* Issuers who participated in HHS–RADV for these benefit years also received issuer-specific and enrollee-specific results in the Audit Tool at the same time the summary information was released. Additionally, HHS conducted two pilot years of HHS–RADV for the 2015 and 2016 benefit years to give HHS and issuers experience with how the audits would be conducted prior to applying HHS–RADV results to adjust issuers' risk scores and risk adjustment transfers in the applicable state market risk pool and for the 2016 benefit year, participating issuers were provided illustrative 2016 benefit year HHS–RADV results based on the application of the current error rate methodology. As noted previously, HHS–RADV was not conducted for the 2014 benefit year so there were no results to release or otherwise share. We also point this commenter to the analysis in the proposed rule,[66] as well as the results of the evaluation of the sliding scale adjustment options in the 2019 RADV White Paper, using 2017 benefit year HHS–RADV results.[67] In addition, Tables 3 and 4 in this rule share an analysis and comparison of the estimated changes reflecting implementation of this policy using both 2017 and 2018 benefit year HHS–RADV results.

3. Negative Error Rate Issuers With Negative Failure Rates

HHS–RADV uses a two-sided outlier identification approach because the long-standing intent has been to account for identified material risk differences between what issuers submitted to their EDGE servers and what was validated in

[64] See, *e.g.,* 85 FR at 33608.

[65] See section 4.4.5 and Appendix C of the 2019 RADV White Paper.

[66] See 85 FR at 33613.

[67] See section 4.4.5 and Appendix C of the 2019 RADV White Paper.

medical records through HHS–RADV, regardless of the direction of those differences.[68] 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 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[69] and the HCC failure rate approach to error estimation finalized in the 2019 Payment Notice.[70]

In response to stakeholder feedback about the impact of negative error rate issuer HHS–RADV adjustments on issuers who are not outliers, we proposed to adopt a constraint to the calculation of negative error rate outlier issuers' error rates in cases when an outlier issuer's failure rate is negative. 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. For this reason, we proposed to refine the error rate calculation to mitigate the impact of adjustments that result from negative error rate outliers that are driven by newly found HCCs rather than by high validation rates.

Beginning with 2019 benefit year HHS–RADV, we proposed to adopt an approach that constrains negative error rate outlier issuers' error rate calculations in cases when an issuer's failure rate is negative. For negative error rate outlier issuers with negative failure rates, the proposed constraint would be applied to the GAF such that this value 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 barred‖ terms and definitions into the error rate calculation[71] process:

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

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

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

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

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

[68] 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.

[69] 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.

[70] 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.

[71] This calculation sequence is expressed here in a revised order compared to how the sequence is published in the 2021 Payment Notice (85 FR at 29196–29198). This change was made to simplify the illustration of how this sequence will be combined with proposals finalized in this rule. The different display does not modify or otherwise change the amendments to the outlier identification process finalized in the 2021 Payment Notice.

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, is the group adjustment factor for HCC failure rate group *G* for an issuer *i*.

We would then compute total adjustments and error rates for each outlier issuer based on the weighted aggregates of the $GAF_{G,i}$.[72]

We are finalizing this refinement to the error rate calculation as proposed. We will adjust the GAF calculation to be the difference between the weighted group mean and zero for negative error rate issuers with negative failure rates beginning with the 2019 benefit year of HHS–RADV.

Comments: Most commenters supported the proposed negative failure rate constraint. These commenters tended to be concerned that the current methodology rewards issuers who fail to submit accurate data to the EDGE server, were concerned about predictability of HHS–RADV adjustments, or thought that the proposed constraint would result in more equitable HHS–RADV adjustments. A few commenters opposed the proposed negative failure rate constraint. These commenters, as well as another commenter that was not opposed to the negative failure rate constraint, expressed concerns that the proposed negative failure rate constraint would treat issuers with different validation rates and the same rate of found HCCs the same for calculating error rates, potentially penalizing issuers that submitted more verifiable HCCs. Some commenters argued that the potential for underreporting of risk in risk adjustment was minor, and one supported allowing issuers to get credit for the risk that they incurred including through newly found HCCs.

Other commenters generally agreed that a change in methodology is needed to reduce the magnitude of HHS–RADV adjustments due to negative error rate issuers and the impact of these adjustments on non-outlier issuers in the same state market risk pool. Some commenters wanted HHS to abandon the two-sided nature of the outlier identification process and not adjust for any negative error rate outliers or urged HHS to look for ways to minimize adverse impact of negative error rate outliers on non-outliers. Other commenters recommended that HHS analyze the failure rates for negative error rate outliers without including found HCCs (meaning that only non-validated EDGE HCCs would be contributing to the issuer's failure rate) and compare the results with the current methodology to assess if negative error rate outliers had better validation rates. Another commenter requested that HHS monitor data on the policy's impact, if finalized.

Response: We are finalizing the proposed approach to constrain negative error rate outlier issuers' error rate calculations in cases when an outlier issuer's failure rate is negative and will apply this constraint beginning with the 2019 benefit year of HHS–RADV. We believe that the negative failure rate constraint to the GAF calculation in the error rate calculation will reduce potential incentives for issuers to use HHS–RADV to identify more HCCs than were reported to their EDGE servers and provide additional incentives for issuers to submit the most accurate data to the EDGE server. It also will mitigate the impact of HHS–RADV adjustments to transfers in the case of negative error rate issuers with negative failure rates and improve predictability. Specifically, this approach would limit the financial impact that negative error rate outliers with negative failure rates will have on other issuers in the same state market risk pool and can be easily implemented under the current error rate methodology.

We understand that this constraint has limitations. We used 2017 and 2018 benefit year HHS–RADV results to analyze the failure rates of negative error rate outliers and explore the impact of excluding found HCCs. We found that negative error rate outliers tended to have better than average validation rates, particularly when the HCC grouping methodology finalized in this rule is applied and those issues get credit for IVA findings that substitute for EDGE HCCs in the same HCC coefficient estimation group. However, at the same time, we recognize that there are limitations to the negative failure rate constraint policy as it does not distinguish between issuers with different validation rates and the same rate of found HCCs. Thus, as previously noted, this policy and the other changes to the error rate calculation in this rule are targeted refinements to the current methodology as we consider other potential long-term approaches. In proposing and finalizing these changes, we sought to balance the goals of promoting stability and predictability of HHS–RADV adjustments and adopting refinements as expeditiously as possible. The negative error rate constraint was designed with these goals in mind, as it builds on the current methodology, which issuers now have several years of experience with, and is easy to implement. It is an interim measure that will limit the financial impact that negative error rate outliers with negative failure rates have on other issuers in the same state market risk

[72] 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_RETIRED_5CR_070519.pdf.*

pool. We remain committed to continuing to explore different longer-term options, including approaches that involve significant methodological changes, such as those described in the 2019 RADV White Paper that would switch to identifying outliers based on risk score instead of number of HCCs.[73]

We also decline to abandon the two-sided nature of the outlier identification process. 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. The increase to risk scores 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[74] and the HCC failure rate approach to error estimation finalized in the 2019 Payment Notice.[75] The two-sided approach also encourages issuers to ensure that their EDGE-reported risk scores reflect the true actuarial risk of their enrollees.

We agree with the commenter that supported allowing issuers to get credit for the risk that they incurred including through newly found HCCs. It ensures that risk adjustment transfers are made based on documented risk and that, consistent with the statute, the HHS-operated program assesses charges to plans with lower-than-average actuarial risk while making payments to plans with higher-than-average actuarial risk. As such, even with the adoption of this constraint, the calculation of error rates will still include found HCCs. The negative failure rate constrained value in the calculation of the GAF will only impact the negative failure rate portion of an issuer's GAF. Therefore, this policy ensures that negative error rate outlier issuers with negative failure rates will only get credit in their error rate calculation for finding HCCs at a similar rate as they reported to EDGE and will not get credit for finding more HCCs in HHS–RADV than they reported on EDGE. We believe that any issuer with a negative failure rate is likely to review their internal processes to better capture missing HCCs in future EDGE data submissions. We intend to monitor the impact of this policy on future benefit years of HHS–RADV data.

Comments: One commenter noted that it is not evident that issuers with negative failure rates in one HCC group are adding more diagnoses given that the three HCC grouping structure allows for HCCs to be found in one grouping and missing in another grouping. One commenter noted that the proposal to calculate the GAF between zero and the weighted mean for negative failure rate issuers does not reflect the outlier portion of the negative error rate outlier (because the adjustment is within the confidence intervals for two of three HCC groupings). Another commenter expressed concerns that the national mean is not adjusted for found HCCs under the proposal leading to concerns that the national mean is being inflated and proposed adjusting negative error rate outliers to the edge of the confidence intervals as an alternative to the proposed negative failure rate constraint.

Response: The purpose of this negative failure rate constraint policy is to mitigate the impact of HHS–RADV adjustments due to negative error rate issuers with negative failure rates. We understand that the HCC failure rate grouping methodology can result in an issuer finding HCCs in one HCC failure rate group when the HCC may be missing in another HCC failure rate grouping. We are finalizing the HCC grouping refinement discussed earlier in this rule to help prevent those cases from occurring when the HCCs are in the same HCC coefficient estimation group in the adult risk adjustment models. We also acknowledge that this constraint would not affect the calculation of the national mean, which would continue to consider all found HCCs and that the calculation of the GAF under this constraint policy may not fully reflect the outlier portion. We considered these limitations and weighted them against the benefits of this policy. While we do have concerns about the impact of adjustments resulting from negative error rate issuers with negative failure rates, we believe that issuers should retain the ability to find HCCs in HHS–RADV. Having the ability to find HCCs in HHS–RADV is important to ensure that issuers' actual actuarial risk is reflected in HHS–RADV, especially when those HCCs replace related HCCs that were reported to EDGE. As such, we believe that found HCCs should continue to contribute to the national mean. At the same time, given the number of negative error rate issuers with negative failure rates, we believe that it is important to refine the current methodology to reduce the incentives for issuers to find HCCs in HHS–RADV that are not reported in EDGE. We intend to monitor the impact of this policy on HHS–RADV adjustments and will continue to explore potential further refinements and changes to the HHS–RADV methodology and program requirements for future benefit years.

Comment: Some commenters stated that the HHS–RADV Protocols and the applicable EDGE data submission requirements did not align and recommended that HHS align these documents. One of these commenters recommended aligning these rules as an alternative to constraining negative error rate outliers with negative failure rates.

Response: We did not propose and are not finalizing any changes to the EDGE data submission requirements. As noted earlier, 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. This includes allowing issuers to get credit for the risk that they incurred including through newly found HCCs. However, in response to stakeholder feedback, we are adopting the negative failure rate constraint to limit the impact of HHS–RADV adjustments due to negative error rate issuers with negative failure rates beginning with the 2019 benefit year of HHS–RADV. We disagree that the HHS–RADV Protocols and the EDGE data submission are not appropriately aligned as the EDGE data submissions and HHS–RADV Protocols are different processes. Specifically, the EDGE data submission process for risk adjustment requires issuers to submit all paid claims to their respective EDGE servers, regardless of provider type, for the applicable benefit year. These paid claims provide the diagnoses that are used to calculate risk adjustment transfers at the state market risk pool level under the state payment transfer formula.[76] HHS–RADV is a review of an enrollee's medical records to confirm the diagnoses used to perform the

[73] See Section 3.3 on addressing the influence of HCC hierarchies on failure rate outlier determination (Pages 63–71). *https://www.cms.gov/files/document/2019-hhs-risk-adjustment-data-validation-hhs-radv-white-paper.pdf.*

[74] 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.

[75] See 83 FR 16930 at 16962.

[76] For the 2014 through 2016 benefit years, EDGE data was also used for the transitional reinsurance program established under section 1341 of the PPACA. The reinsurance program provided reimbursement based on the total amount of claims paid. Beginning with the 2018 benefit year, EDGE data is also used for calculating payments under the high-cost risk pool (HCRP) parameters added to the HHS risk adjustment methodology. Similar to the reinsurance program, HCRP payments are based on the amount of paid claims. Therefore, information on all claims paid—from all provider types—for a given benefit year should be submitted by issuers to their EDGE servers.

calculations under the state payment transfer formula. HHS- RADV allows issuers to take into account an issuer's paid claims for the applicable benefit year for medical record review and this process also allows issuers to take into account certain diagnoses found during the review of the medical records of the enrollee to provide a more complete and accurate picture of an enrollee's risk to the issuer. Further, while HHS–RADV Protocols allow IVA and SVA auditors to abstract documented ''Lifelong Permanent Conditions'' [77] that may not be captured in EDGE data submissions, we disagree that such an approach is inappropriate. The list of Lifelong Permanent Conditions is a set of health conditions that require ongoing medical attention and where all associated diagnoses are typically unresolved once diagnosed. Allowing abstraction of diagnosis codes for those conditions from medical records submitted during HHS–RADV if the Lifelong Permanent Condition is identified in the enrollee's medical history included in a medical record for the applicable benefit year ensures that an enrollee's full health risk is captured and reflected in risk adjustment transfers for that state market risk pool.

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

As discussed elsewhere in this final rule, we are finalizing as proposed each of the three constituent proposals to refine the current error rate calculation. To illustrate the interaction of the finalized policies to create Super HCCs for HHS–RADV grouping purposes, apply the sliding scale adjustment, and constrain negative failure rates for negative error rate outliers, this section outlines the complete finalized revised error rate calculation methodology formulas that will apply beginning with the 2019 benefit year of HHS–RADV, integrating all the changes finalized in this rule.[78]

First, HHS will 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 finalized approach, Super HCCs will be defined as HCCs that have been aggregated such that HCCs that are in the same HCC coefficient estimation group in the adult models are aggregated together and all other HCCs each compose a Super HCC individually. Using the Super HCCs, we will calculate the HCC failure rate as follows:

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

Where:

c is the index of the cth 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, will be grouped according to the current sorting algorithm in the current HHS–RADV failure rate grouping methodology.[79] These HCC groupings will 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, thereby assigning each Super HCC into a high, medium, or low HCC failure rate grouping. This process ensures that all HCCs in a Super HCC are grouped into the same HCC failure rate 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 (or SVA, 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:

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

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

[77] See, for example, Appendix E of the 2018 Benefit Year HHS–RADV Protocols, which describes the guidelines for abstracting Lifelong Permanent Conditions from medical records for purposes of the 2018 benefit year of HHS–RADV.

[78] The illustration of the error rate calculation methodology formulas that will apply beginning with the 2019 benefit year of HHS–RADV also includes the policy finalized in the 2021 Payment Notice to not consider issuers with fewer than 30 HCCs in an HCC failure rate group to be outliers in that HCC failure rate group but continue to include such issuers in the calculation of national metrics. See 85 FR at 29196–29198.

[79] See Section 11.3.1 of the 2018 HHS–RADV Protocols at *https://www.regtap.info/uploads/library/HRADV_2018Protocols_070319_RETIRED_5CR_070519.pdf* for a description of the process prior to the introduction of Super HCCs. Beginning with the 2019 benefit year of HHS–RADV, Super HCCs would take the place of HCCs in the process. The 2019 HHS–RADV Protocols have thus far only been published in part at *https://www.regtap.info/uploads/library/HRADV_2019_Protocols_111120_5CR_111120.pdf.* The section of the 2019 HHS–RADV Protocols pertaining to HCC grouping for failure rate calculations is not included in the current version. Once published, this section will be updated to include steps related to creation of Super HCCs.

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 will then be compared to the national metrics for each HCC failure rate grouping. If an issuer's failure rate for an HCC failure rate 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 failure rate group, the failure rate for the issuer's HCCs in that group will be considered an outlier (if the issuer meets the minimum number of HCCs for the HCC failure rate group). Based on issuers' failure rates for each HCC failure rate group, outlier status will 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 EDGE HCCs. If no issuers' HCC group failure rates in a state market risk pool materially deviate from the national mean of failure rates or if those issuers whose failure rates do materially deviate from the national mean do not also meet the minimum HCC frequency requirement (that is, if no issuers in the state market risk pool are outliers), HHS will not apply any HHS–RADV adjustments to issuers' risk scores or to transfers in that state market risk pool.

Then, once the outlier issuers are determined, we will calculate the GAF taking into consideration the outlier issuer's distance from the confidence interval and limiting calculation of the GAF when if the issuer is a negative error rate outlier with a negative failure rate. The formula [80] will apply as follows:

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

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

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

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

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

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 $freqEDGE_{G,i} < 30$ 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:

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

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

• $disZ_{G,i,r}$ is the z-score of issuer i's $GFR_{G,i}$, for HCC failure rate group G discounted according to the sliding scale adjustment 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' GFRs for HCC failure rate group G;

• $\mu\{GFR_G\}$ is the weighted national mean of all issuers' GFRs for HCC failure rate group G.

Once an outlier issuer's GAF is calculated, the enrollee adjustment will be calculated by applying the GAF to an enrollee's individual EDGE HCCs. For example, if an issuer has an 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 EDGE HCC risk score coefficients 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 strata 1 through 9 with EDGE HCCs: [81]

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

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 EDGE HCCs in strata 1 through 9.

Next, for each sampled enrollee with EDGE HCCs, HHS will 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 IVA or SVA 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 with EDGE HCCs.

After calculating the outlier issuers' sample enrollees with HCCs' adjusted EDGE risk scores, HHS will calculate an outlier issuer's error rate by extrapolating the difference between the amended risk score and EDGE risk score for all enrollees (strata 1 through 10) in the sample. The extrapolation formula will 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 EDGE HCCs will 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 will be:

[81] Some enrollees sampled in Strata 1–3 will only have RXCs, which are not considered as part of the determination of an enrollee adjustment factor.

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

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 will apply the outlier issuer's error rate to adjust that issuer's applicable benefit year's plan liability risk score.[82] This risk score change, which also will impact the state market average risk score, will then be used to adjust the applicable benefit year's risk adjustment transfers for the applicable state market risk pool.[83] Due to the budget-neutral nature of the HHS-operated risk adjustment program, adjustments to one issuer's risk scores and risk adjustment transfers based on HHS–RADV findings affect 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.

In the proposed rule, 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. We performed a similar analysis using 2018 benefit year HHS–RADV results, once the data became available. Table 3 provides a comparison of the national failure rate metrics under the current and new, finalized methodologies using 2017 and 2018 benefit year HHS–RADV results. Additionally, using the 2017 and 2018 HHS–RADV data, Table 4 provides a comparison between the estimated mean error rates using the current methodology for sorting HCCs for HHS–RADV grouping or the finalized Super HCC aggregation for sorting of HCCs for HHS–RADV groupings, with the finalized negative failure rate constraint and the finalized sliding scale adjustment also being applied. As shown in Tables 3 and 4, the analysis of 2018 HHS–RADV results provided roughly the same figures as the 2017 HHS–RADV results, and offers further support for finalizing these refinements to the error rate calculation.

TABLE 3—A COMPARISON OF HHS–RADV NATIONAL FAILURE RATE METRICS BASED ON PRIOR BENEFIT YEAR HHS–RADV DATA

HHS–RADV data benefit year	Group	Weighted mean failure rate		Weighted std. dev.		Lower threshold		Upper threshold	
		Current grouping	New grouping	Current grouping	New grouping	Current grouping and 95% CI	New grouping and 90% CI	Current grouping and 95% CI	New grouping and 90% CI
2017 Data	Low	0.0476	0.0496	0.0973	0.0959	−0.1431	−0.1082	0.2382	0.2074
	Med	0.1549	0.1557	0.0992	0.0994	−0.0395	−0.0078	0.3493	0.3192
	High	0.2621	0.2595	0.1064	0.1065	0.0536	0.0843	0.4706	0.4347
2018 Data	Low	0.0337	0.0369	0.0884	0.0856	−0.1396	−0.1038	0.2070	0.1777
	Med	0.1198	0.1225	0.0862	0.0856	−0.0490	−0.0184	0.2887	0.2633
	High	0.2262	0.2283	0.0919	0.0914	0.0461	0.0779	0.4062	0.3787

TABLE 4—A COMPARISON OF HHS–RADV ERROR RATE (ER) ESTIMATED CHANGES BASED ON PRIOR BENEFIT YEAR[84] HHS–RADV DATA

Scenario	2017 Data				2018 Data			
	Current sorting method		New sorting method		Current sorting method		New sorting method	
	Mean neg. ER (%)	Mean pos. ER (%)	Mean neg. ER (%)	Mean pos. ER (%)	Mean neg. ER (%)	Mean pos. ER (%)	Mean neg. ER (%)	Mean pos. ER (%)
Sorting Method Only	−5.68	9.96	−5.98	9.91	−6.92	5.43	−7.06	5.71
Sorting Method with Negative Constraint	−3.11	9.96	−3.38	9.91	−3.35	5.43	−3.16	5.89
Sorting Method with Sliding Scale[85]	−2.27	5.28	−2.49	5.32	−3.07	2.21	−3.21	2.45
Sorting Method, Sliding Scale & Negative Constraint (Finalized)	−1.50	5.28	−1.66	5.32	−1.71	2.21	−1.86	2.47

[82] 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. As detailed in Section II.B, we are finalizing the transition to the concurrent application of HHS–RADV results such that issuer risk score error rates for non-exiting issuers will also be applied to the risk scores and transfer amounts for the benefit year being audited beginning with the 2020 benefit year of HHS–RADV.

[83] See 45 CFR 153.350(c).

[84] These estimates reflect the exclusion from outlier status of those 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 or 2018 benefit year HHS–RADV. We excluded issuers with fewer than 30 HCCs 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.

[85] This analysis reflects the sliding scale policy finalized in Section II.A.2. of this rule which creates a sliding scale adjustment from +/−1.645 to 3 standard deviations.

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.[86] 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.[87] [88]

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 timeline for the reporting, collection, and disbursement of HHS–RADV adjustments to transfers [89] and the changes to the risk adjustment holdback policy,[90] both of which lead to reopening of prior year risk adjustment transfers, we proposed to switch away from the prospective approach for non-exiting issuers. We proposed to make the transition and apply HHS–RADV results to the benefit year being audited for all issuers starting with the 2021 benefit year of HHS–RADV. We proposed applying HHS–RADV results to the benefit year being audited for all issuers in an effort 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.

In the proposed rule, we explained that if we finalized and implemented 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. Then, 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).[91]

In order to effectuate this transition, we proposed 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 plan liability risk scores and risk adjustment transfers.[92] This approach would result in one final HHS–RADV adjustment to 2021 benefit year plan liability risk scores and risk adjustment 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.[93]

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 were taken into consideration in calculating risk adjustment plan liability risk scores, we also proposed an alternative approach: the ''combined plan liability risk score option.'' 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 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. We sought 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.

We also explained in the proposed rule that the transition to a policy to apply HHS–RADV results to the benefit year being audited for all issuers would remove the need to continue the current policy on issuers entering sole issuer markets finalized in the 2020 Payment Notice.[94] 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.[95] We further explained that 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 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 if there was an outlier issuer in the same state market risk pool.

We also sought comment on an alternative timeline, in which HHS 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. We explained that under the alternative timeframe, 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. Lastly, we sought

[86] See 78 FR 15410 at 15438.

[87] 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.*

[88] 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 plan liability risk scores, resulting in adjustments to 2017 benefit year risk adjustment transfer amounts. See 83 FR at 16965–16966. We updated this policy to only apply HHS–RADV results for exiting issuers that are positive error rate outliers beginning with the 2018 benefit year. See the 2020 Payment Notice, 84 FR at 17503–17504.

[89] See 84 FR at 17504 through 17508.

[90] 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.*

[91] 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.

[92] See Section 5.2 of the 2019 RADV White Paper.

[93] 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.

[94] 84 FR at 17504.

[95] Ibid.

comment on whether, if we finalized and implemented either of the transition options using the alternative timeline, we should also pilot RXCs for the 2020 benefit year HHS–RADV.

We are finalizing the proposed transition from the current prospective application of HHS–RADV results for non-exiting issuers and will apply HHS–RADV audit findings to the benefit year being audited for all issuers, starting with the 2020 benefit year HHS–RADV, by combining 2019 and 2020 benefit years HHS–RADV results for non-exiting issuers following the average error rate approach. We also reaffirm that, as a result of finalizing these changes, we will not need to continue the current policy on issuers entering sole issuer markets after the transition is effectuated. Therefore, if a new issuer entered a state market risk pool in 2020, its risk adjustment plan liability risk score(s) and transfer for 2020 benefit year risk adjustment could be impacted by the new issuer's own 2020 HHS–RADV results and the combined 2019 and 2020 HHS–RADV results of other issuers in the same state market risk pool. For exiting issuers, HHS will continue to adjust only for positive error rate outliers, as opposed to both positive and negative error rate outliers.[96] Beginning with the 2021 benefit year of HHS–RADV, plan liability risk scores and risk adjustment transfers will only be adjusted once based on the same benefit year's HHS–RADV results (that is, 2021 benefit year HHS–RADV results would adjust 2021 benefit year plan liability risk scores and transfers for all issuers).[97] Additionally, HHS will continue to pilot RXCs for the 2020 benefit year.

We are finalizing this change to apply HHS–RADV results to the benefit year being audited for all issuers 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. In addition, this 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. Once the transition is effectuated, it will also prevent situations in which an issuer who newly enters a state market risk pool, including new market entrants to a sole issuer market, is subject to HHS–RADV adjustments from the prior benefit year for which they did not participate.

Comments: The majority of commenters supported switching from the prospective application of the HHS–RADV results to the benefit year being audited. These commenters generally agreed that having a concurrent application would maintain actuarial soundness in the application of an issuer's HHS–RADV error rate, provide stability to HHS–RADV results, and promote fairness in the HHS–RADV process. One commenter suggested that HHS should consider maintaining the current prospective application of HHS–RADV findings; another commenter suggested HHS exempt new issuers from having their transfers adjusted due to HHS–RADV.

Regarding the transition year, some commenters supported switching to the concurrent application in the 2021 benefit year as proposed due to concerns that changing the transition year to the 2020 benefit year of HHS–RADV would heighten the already significant uncertainty surrounding 2020 as a result of COVID–19, with one commenter noting that issuers did not account for this change in their 2020 pricing. However, most commenters supported switching to the concurrent application with the 2020 benefit year, suggesting that it would be most appropriate to transition to a concurrent application as early as possible and one cited to the various changes to the HHS-operated risk adjustment program beginning with the 2021 benefit year as further support for the alternative timeline for the transition. One commenter requested additional information on the 2020 benefit year HHS–RADV timeline.

Response: We are finalizing the proposal to switch from the current prospective application of the HHS–RADV results to the benefit year being audited, starting with the 2020 benefit year. As previously noted, when we finalized the prospective HHS–RADV results application policy, we did not anticipate the extent of changes that could occur in the risk profile of enrollees or market participation by issuers from benefit year to benefit year. As a result of experience over the early years of the program, we believe that transitioning to apply HHS–RADV results on a concurrent basis for all issuers will provide greater stability, promote fairness, and enhance actuarial soundness, specifically in the event that an issuer's risk profile, enrollment, or market participation changes significantly from benefit year to benefit year. In light of the other changes to HHS–RADV program operations described in this rule which will lead to reopening of prior benefit year risk adjustment transfers,[98] it is also no longer necessary to apply HHS–RADV results on a prospective basis to allow time to complete the discrepancy and appeals processes to avoid having to reopen prior year transfers. We also agree that we should begin the application of the results on a concurrent basis as soon as possible and will implement the policy starting with the 2020 benefit year. We believe that starting with the 2020 benefit year will add stability in the midst of the COVID–19 pandemic, as the results from the 2019 and 2020 benefit years of HHS–RADV will be averaged together to calculate the adjustment to 2020 benefit year risk adjustment risk scores. We believe this added stability will account for concerns that issuers did not take this proposed change into consideration when setting rates for the 2020 benefit year. We also agree with the commenter who cited the risk adjustment program updates that apply beginning with the 2021 benefit year as further support for effectuating the transition beginning with the 2020 benefit year.[99]

We did not propose and are not finalizing a new exemption from HHS–RADV for new market entrants. The inclusion of new market entrants in HHS–RADV ensures that those issuers' actuarial risk for the applicable benefit year is accurately reflected in risk adjustment transfers, and that the HHS-operated risk adjustment program assesses charges to plans with lower-than-average actuarial risk while making payments to plans with higher-than-average actuarial risk. However, new market entrants will no longer be impacted by a prior year's HHS–RADV results and will only be impacted by the results from the benefit year under which they participated in the state market risk pool after the transition is effectuated.[100]

[96] In addition, positive error rate outlier issuers' 2019 and 2020 HHS–RADV results will be applied to the risk scores and transfers for the benefit year being audited. The average error rate approach is not applicable because exiting issuers who participated in 2019 HHS–RADV would not have 2020 benefit year risk scores or transfers to adjust.

[97] 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.

[98] Ibid.

[99] For example, in the 2021 Payment Notice, we finalized several updates to the HHS–HCC clinical classification to develop updated risk factors that apply beginning with the 2021 benefit year risk adjustment models. See 85 FR at 29175.

[100] As noted above, a new entrant to a state market risk pool in 2020 would see its risk score(s) and transfer impacted by the new issuer's own 2020 HHS–RADV results, the combined 2019 and 2020 HHS–RADV results of other non-exiting issuers in the same state market risk pool, and the 2020 HHS–RADV results for positive error rate outlier exiting

Continued

HHS intends to provide more information on the 2020 benefit year HHS–RADV timeline in the future, but generally anticipates it will commence as usual with the release of samples in May 2021. As previously noted in this rule, HHS has provided details on the updated timeline on the activities for 2019 benefit year HHS–RADV.[101]

Comments: Most commenters who submitted comments on the options for combining HHS–RADV results during the transition period supported using the average error rate approach, noting that it would provide more stability and transparency than the combined plan liability risk score option. One commenter who expressed a preference for the average error rate approach cited concerns with the amplifying effect of adjusting risk scores twice under the plan liability risk score option. Most commenters who supported the average error rate approach supported effectuating the transition using 2019 and 2020 benefit years' error rate results. These commenters noted that aggregating the results of these 2 years could reduce volatility and smooth over potential challenges issuers may face when conducting HHS–RADV audits for these benefit years due to the COVID–19 public health emergency. A few commenters who supported use of the average error rate approach urged HHS to implement the transition and use 2020 and 2021 benefit years' results, suggesting it would be the most straightforward approach. One commenter requested clarification as to whether the average error rate approach would use a weighted average error rate.

A few commenters supported the combined plan liability risk score option for the transition years of HHS–RADV. One of these commenters believed that the combined plan liability risk score option would be a fairer way to provide consistency, while a different commenter that supported the combined plan liability risk score option was concerned that the average error rate approach would reduce the otherwise applicable HHS–RADV adjustment. Another commenter compared the two alternative approaches, noting that the average error rate would align well with some issuers' practices, while the combined liability risk score option would align better with other issuers' financial reporting.

Response: We are finalizing the use of the average error rate approach to transition to the concurrent application of HHS–RADV results for non-exiting issuers by combining their 2019 and 2020 benefit years' HHS–RADV results. In response to comments we clarify that for simplification purposes, HHS will apply an unweighted average value of the 2019 and 2020 benefit years' HHS–RADV results to adjust 2020 benefit year risk scores and transfers. We proposed using a combined plan liability risk score as an alternative option, believing that it could provide a more consistent transition to a concurrent application of HHS–RADV results. However, the majority of comments on these transition options emphasized the extent to which they believed an average error rate approach will actually provide greater stability and transparency for the HHS–RADV adjustments applied during the transition period. After consideration of comments, we agree that the average error rate approach will be the optimal transitional approach. More specifically, aggregating the 2019 and 2020 benefit years' results for non-exiting issuers and using the unweighted average value of those benefit years' HHS–RADV results to adjust transfers will allow for more consistency, reduce potential volatility, and better accommodate any potential disparities or challenges due to COVID–19. As noted previously, we also believe the transition to the application of the results on a concurrent basis should be implemented as soon as possible and therefore will start the concurrent application of HHS–RADV results for all issuers starting with the 2020 benefit year. We recognize that there are advantages to the combined plan liability risk score option, which is why we proposed it for combining HHS–RADV results for the transition years. However, for the reasons outlined above, we believe the average error rate method is the more balanced approach to effectuate the transition and combine 2019 and 2020 HHS–RADV results for non-exiting issuers.

Comments: Some commenters suggested HHS cancel either the 2019 or 2020 benefit years of HHS–RADV. One of these commenters expressed concern that the COVID–19 pandemic could potentially skew the 2020 benefit year HHS–RADV results. Other commenters stated that COVID–19 would make it difficult for providers to respond to issuer requests for the medical documentation needed to complete audits, which they noted could skew HHS–RADV results.

Response: We appreciate the concerns related to the potential impact of COVID–19, but are not cancelling HHS–RADV for either the 2019 or 2020 benefit year. We believe that cancelling either year of this program would be detrimental to program integrity and would result in future difficulties monitoring HHS–RADV trends. We acknowledge that the COVID–19 pandemic puts a number of stressors on providers and issuers. Recognizing the impact of the public health emergency on HHS–RADV activities, we postponed the start of 2019 benefit year HHS–RADV activities.[102] As recently announced, IVA samples for 2019 benefit year HHS–RADV will be released in January 2021 and we anticipate 2020 benefit year HHS–RADV will commence as usual.[103] We will continue to monitor the COVID–19 public health emergency and will consider whether additional flexibilities for HHS–RADV are appropriate. Further, as noted above, the adoption of the average error rate approach for the transition to the concurrent application of HHS–RADV is intended to help reduce volatility related to potential challenges issuers may face when conducting HHS–RADV audits for these benefit years due to the COVID–19 public health emergency.

Comments: Most commenters supported continuing the pilot of RXCs for the 2020 benefit year. Some of these commenters suggested that continuing to pilot RXCs would allow for more consistency between 2019 and 2020 and support transitioning to the concurrent application of HHS–RADV results starting with the 2020 benefit year, while another commenter believed that it would minimize the amount of changes occurring at once. One commenter noted that extending the RXC pilot would benefit the issuers who are still learning how to conduct HHS–RADV for RXCs. Another commenter did not believe it would be necessary to continue piloting RXCs in 2020, but acknowledged that an additional pilot period would allow issuers to focus on HHS–RADV during the COVID–19 pandemic, rather than adjusting to new aspects of HHS–RADV reporting.

Response: After consideration of comments, we are finalizing the continuation of the pilot for RXCs for the 2020 benefit year. Extending the RXC pilot an additional benefit year will increase consistency between the

issuers in the same state market risk pool. However, a new entrant to a state market risk pool in 2021 would see its risk score(s) and transfer impacted by 2021 HHS–RADV results only.

[101] See the ''2019 Benefit Year HHS–RADV Activities Timeline'' *https://www.regtap.info/uploads/library/HRADV_Timeline_091020_5CR_091020.pdf.*

[102] *https://www.cms.gov/files/document/2019-HHS-RADV-Postponement-Memo.pdf.*

[103] See the ''2019 Benefit Year HHS–RADV Activities Timeline'' *https://www.regtap.info/uploads/library/HRADV_Timeline_091020_5CR_091020.pdf.*

operations of the 2019 and 2020 benefit years' HHS–RADV and facilitate the combination of the HHS–RADV adjustments for these benefit years as we transition to a concurrent application of HHS–RADV results starting with the 2020 benefit year. We agree with commenters who suggested that an additional pilot year for RXCs would benefit issuers and provide an opportunity to continue to improve their internal process for conducting HHS–RADV for RXCs.

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 final rule, we are finalizing the modifications to the calculation of error rates to modify the HCC failure rate grouping methodology for HCCs that share an HCC coefficient estimation group in the adult risk adjustment models; to calculate and apply a sliding scale adjustment for cases where outlier issuers are near the confidence intervals; and to constrain the error rate calculation for issuers with negative failure rates. We are also finalizing the transition from the current prospective application of HHS–RADV results [104] to apply the results to the benefit year being audited. These are methodological changes to the error estimation 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 did 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.[105]

IV. Regulatory Impact Statement

A. Statement of Need

This rule finalizes standards related to HHS–RADV, 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 earlier provided detail on the implementation of HHS–RADV.

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 final rule, we considered numerous alternatives to the presented policies. Below we discuss the key regulatory alternatives considered.

We considered an alternative approach to the 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 finalized approach that focuses on HCCs that share an HCC coefficient estimation group and have 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.[106]

We considered several other options for addressing the payment cliff effect besides the specific sliding scale adjustment that we are finalizing. 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.[107] 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 less stability, since the vast majority of outlier issuers had their original failure rates applied without the benefit of subtracting the weighted mean difference.[108] In addition, while the original methodology did not specifically result in a payment cliff effect, it would have resulted in more and larger adjustments to transfers.

The second option we considered to mitigate the impact of the payment cliff 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,[109] we had 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

[104] The exception to the current prospective application of HHS–RADV results is for exiting issuers identified as positive error rate outliers, whose HHS–RADV results are applied to the risk scores and transfer amounts for the benefit year being audited.

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

[106] See 83 FR 16961 and 16965.

[107] See 79 FR 13755–13770.

[108] See the 2019 RADV White Paper at pages 78–79 and Appendix B.

[109] See 84 FR 17507–17508. See also the 2019 RADV White Paper at page 80.

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 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.[110] In addition, the two-sided outlier identification, and the resulting adjustments to outlier issuer risk scores that 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 during HHS–RADV. 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 to mitigate 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 99.7 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 99.7 percent confidence interval (beyond three standard deviations), and the current significant adjustment to the HCC group weighted mean after three standard deviations. Commenters supported this sliding scale option because it addressed the payment cliff issue without increasing the number of issuers identified as outliers. However, while we recognized that this alternative also would mitigate the payment cliff effect, it would weaken HHS–RADV 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 this rule, we described two of those options. The third option is the ''RA transfer option.'' The RA transfer option would separately calculate 2019 benefit year HHS–RADV adjustments to 2020 benefit year transfers and 2020 benefit year HHS–RADV adjustments to 2020 benefit year transfers.[111] Under this option, we would then calculate the difference between each of these values and the unadjusted 2020 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 2020 benefit year transfers. That is, HHS would separately calculate adjustments for the 2019 and 2020 benefit year HHS–RADV results and incorporate 2019 and 2020 benefit year HHS–RADV results in one final adjustment to 2020 benefit year transfers that would be collected and paid in accordance with the 2020 benefit year HHS–RADV timeline.[112] 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 being finalized.

V. Regulatory Flexibility Act

The RFA (5 U.S.C. 601 *et seq.*) 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 (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 final rule, we establish standards for HHS–RADV. 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.[113] 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 [114] submissions for the 2017 MLR reporting year, 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 604 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 final rule would not affect small rural hospitals. Therefore, the Secretary has determined that this final

[110] 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.

[111] See section 5.2 of the 2019 RADV White Paper.

[112] 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.

[113] *https://www.sba.gov/document/support--table-size-standards.*

[114] Available at *https://www.cms.gov/CCIIO/Resources/Data-Resources/mlr.html.*

rule will not have a significant impact on the operations of a substantial number of small rural hospitals.

VI. 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 2020, that threshold is approximately $156 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.

VII. 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 final rule, we attempted to balance the states' interests in regulating health insurance issuers with the need to ensure market stability and adopt refinements to HHS–RADV standards. 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 final 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.

VIII. 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 final rule is not subject to the requirements of Executive Order 13771 because it is expected to result in no more than de minimis costs.

IX. Conclusion

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

Dated: November 18, 2020.

Seema Verma,

Administrator, Centers for Medicare & Medicaid Services.

Dated: November 23, 2020.

Alex M. Azar II,

Secretary, Department of Health and Human Services.

[FR Doc. 2020–26338 Filed 11–25–20; 4:15 pm]

BILLING CODE 4120–01–P

DEPARTMENT OF COMMERCE

National Oceanic and Atmospheric Administration

50 CFR Part 635

[Docket No.: 201124–0317]

RTID 0648–XT038

Atlantic Highly Migratory Species; 2021 Atlantic Shark Commercial Fishing Year

AGENCY: National Marine Fisheries Service (NMFS), National Oceanic and Atmospheric Administration (NOAA), Commerce.

ACTION: Final rule; fishing season notification.

SUMMARY: This final rule establishes the 2021 opening date for all Atlantic shark fisheries, including the fisheries in the Gulf of Mexico and Caribbean. This final rule also establishes the shark fisheries quotas for the 2021 fishing year, with adjustments based on harvest levels during 2020, and establishes the large coastal shark (LCS) retention limits for directed shark limited access permit holders. NMFS may increase or decrease these retention limits for directed shark limited access permit holders during the year, in accordance with existing regulations, to provide equitable fishing opportunities for commercial shark fishermen in all regions and areas, to the extent practicable. These actions could affect fishing opportunities for commercial shark fishermen in the northwestern Atlantic Ocean, including the Gulf of Mexico and Caribbean Sea.

DATES: This rule is effective on January 1, 2021. The 2021 Atlantic commercial shark fishing year opening dates and quotas are provided in Table 1 under **SUPPLEMENTARY INFORMATION**.

ADDRESSES: Atlantic Highly Migratory Species (HMS) Management Division, 1315 East-West Highway, Silver Spring, MD 20910; *https://www.fisheries.noaa.gov/topic/atlantic-highly-migratory-species.*

FOR FURTHER INFORMATION CONTACT: Lauren Latchford (*lauren.latchford@noaa.gov*), Guy Eroh (*guy.eroh@noaa.gov*), or Karyl Brewster-Geisz (*karyl.brewster-geisz@noaa.gov*) at 301–427–8503.

SUPPLEMENTARY INFORMATION:

Background

The Atlantic commercial shark fisheries are managed primarily under the authority of the Magnuson-Stevens Fishery Conservation and Management Act (Magnuson-Stevens Act). The 2006 Consolidated Atlantic HMS Fishery Management Plan (FMP) and its amendments are implemented by regulations at 50 CFR part 635. For the Atlantic commercial shark fisheries, the 2006 Consolidated HMS FMP and its amendments established, among other things, measures related to commercial shark retention limits, commercial quotas for species and management groups, and accounting for under- and overharvests in the shark fisheries. Regulations include adaptive management measures, such as flexibility in establishing opening dates for the fishing season and the ability to make inseason trip limit adjustments, which provide management flexibility in furtherance of equitable fishing opportunities, to the extent practicable, for commercial shark fishermen in all regions and areas.

On September 29, 2020, NMFS published a proposed rule (85 FR 60947) regarding management measures for the commercial shark fisheries for the 2021 fishing year. The rule proposed opening all Atlantic commercial shark management groups on January 1, 2021, setting initial retention limits for large coastal sharks (LCS) by directed shark