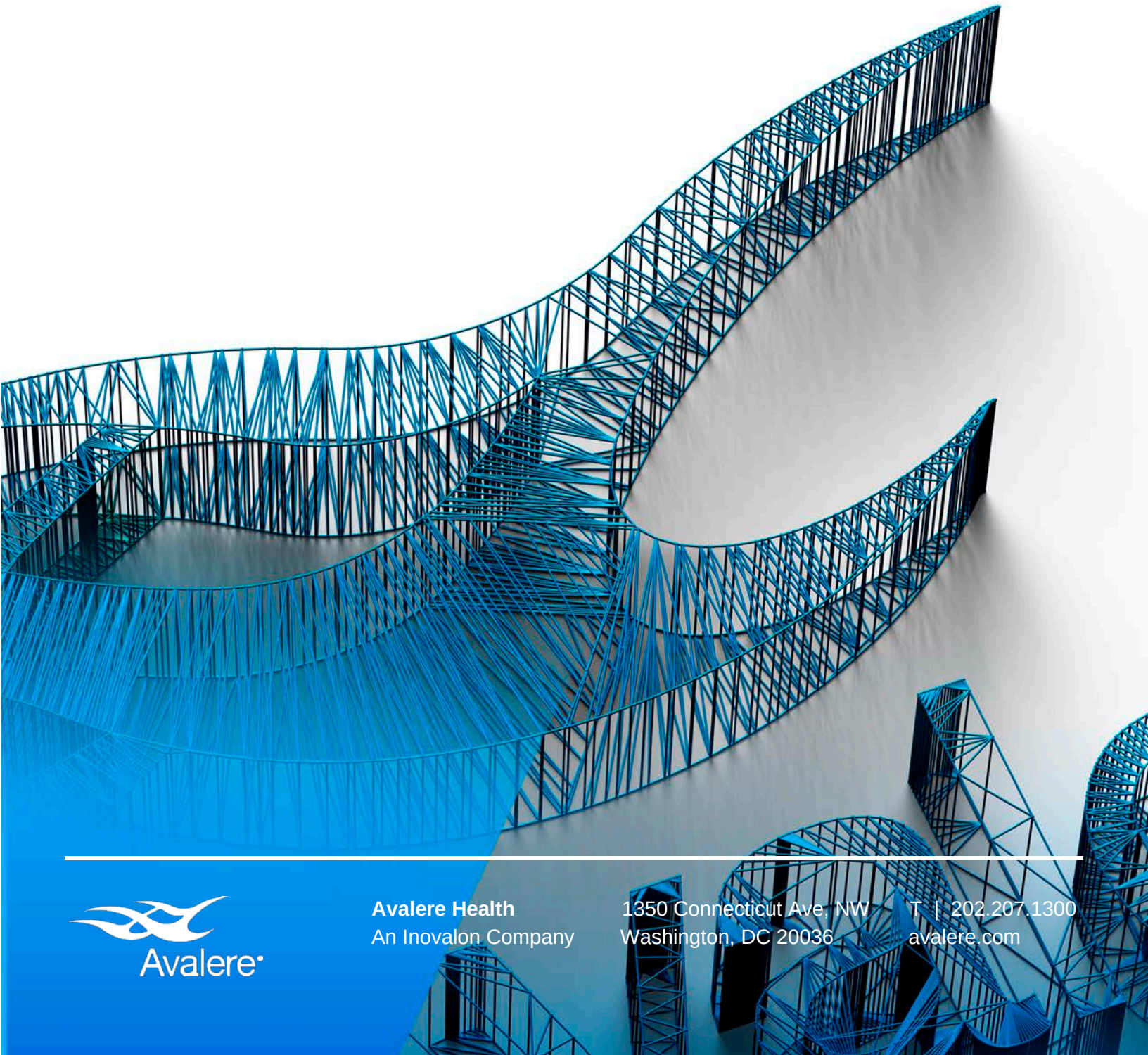
Impact of Eliminating the FFS Adjuster May Vary Based on Plan Enrollee Characteristics



Avalere Health
An Inovalon Company

1350 Connecticut Ave, NW
Washington, DC 20036

T | 202.207.1300
avalere.com

Table of Contents

Funding for this research was provided by Cigna. Avalere Health retained full editorial control.

Avalere

Executive Summary

This study builds upon a March 2019 Avalere paper ("Eliminating the FFS Adjuster from the RADV Methodology May Affect Plan Payment"), analyzing the impact that coding and documentation discrepancies in fee-for-service (FFS) Medicare may have on payments to Medicare Advantage (MA) plans. The prior Avalere paper evaluated a study published by the Centers for Medicare & Medicaid Services (CMS) in October 2018 and found that CMS significantly underestimated the impact that FFS coding discrepancies have on MA payments. As discussed below, new analysis indicates that the payment impact associated with FFS coding discrepancies is greater for certain subgroups of beneficiaries enrolled in the MA program, including dual-eligible beneficiaries and beneficiaries with certain common and potentially serious health conditions (e.g., diabetes, congestive heart failure (CHF)). Given the statutory requirement that MA payments be appropriately adjusted to account for the needs of vulnerable subgroups, our findings raise policy questions that warrant further consideration.[1]

February 2012 Risk Adjustment Data Validation Policy

In February 2012, CMS announced a final payment error calculation methodology for its contract-level Risk Adjustment Data Validation (RADV) audits of MA plans. The final RADV methodology included an offset, the FFS Adjuster, calculated based on a review of medical records submitted to support FFS claims. Using the FFS Adjuster, CMS intended to account for the different documentation standard used to develop the MA risk adjustment model compared to the documentation standards in FFS. CMS did not publicly release a specific estimate for the FFS Adjuster, but estimated that diagnosis coding errors in the FFS claims data had a negative impact on MA payment risk scores of between 4.8% and 8.1%.[2]

Recent Proposals and Actions to Modify the RADV Policy

In a November 2018 Proposed Rule, CMS proposed to revise its RADV methodology to exclude the FFS Adjuster in its payment recoupment calculations. CMS cited an internal study published by the Center for Program Integrity in October 2018 (CPI Study). According to CMS, the CPI Study showed that FFS coding discrepancies do not have a meaningful impact on risk adjustment model estimation and, further, do not bias MA plan payment. As a result, CMS concluded that no FFS Adjuster is necessary.

1 42 USC 1395w-23(a)(C)(1)(i) (health status adjustments in general); 1395w-23(a)(1)(C)(iii) (special needs individuals); 1395w-23(a)(1)(I)(III) (dual eligible individuals); 1395w-23(a)(1)(I)(IV)-(V) (mental health, substance use, chronic kidney disease, and end-stage renal disease). Notably, CMS recently requested public comment as to whether these provisions "mandate[] an FFS Adjuster, prohibit[] an FFS Adjuster, or should otherwise be read to inform our proposal not to apply an FFS Adjuster." 84 Fed. Reg. 30983 (June 28, 2019).

2 CMS recently released information indicating the existence of a subsequent analysis of the need for the FFS Adjuster conducted in March 2018. While publicly available information about this study is limited, it appears to have been modified and re-run as part of the internal study that CMS published in October 2018.

In the same rule, CMS also proposed to alter its contract level RADV methodology to include audits of "subcohorts" of enrollees, such as enrollees who have been diagnosed with a particular condition. More recently, in April 2019, CMS announced that the Contract Year 2014 (CON14) audits would be conducted on a subcohort basis, with a primary focus on beneficiaries diagnosed with diabetes mellitus. CMS also announced that it would audit an additional sample of 299 enrollees drawn from across all contracts who were predicted to have high payment error, further modifying the established RADV sampling strategy.

In addition, the Office of the Inspector General within the Department of Health and Human Services (HHS-OIG), which has previously conducted its audits at the contract-level and deferred to CMS on issues surrounding extrapolation, also recently announced subcohort audits. For example, an audit reported in April 2019 focused solely on diagnoses of acute stroke and major depressive disorder.[3]

March 2019 Avalere Analysis of Impact of FFS Discrepancies on MA Payments

Our prior paper found that certain key assumptions embedded in CMS' internal study do not appropriately capture the full variation in the data and minimize the impact of FFS documentation error on MA payments. Avalere tested alternatives to the assumptions made by CMS and found that, even otherwise accepting the methodology used in the CPI Study,[4] audit miscalibration bias may yield underpayments to MA plans of nearly 8% (see Table 1).[5, 6] In particular, Avalere's analysis found that a root cause of CMS' underestimation of the payment impact was CMS' use of the average number of claims in calculating hierarchical condition category (HCC) person-level error rates. Due to the skewed distribution of claim counts, using the average number severely underestimates the error rate.

Follow-on Analysis of Variation of Payment Impact of FFS Coding Discrepancies by Select Subpopulations

Building on its original study, in this paper Avalere examines the differential payment impact associated with FFS coding discrepancies based on beneficiary eligibility status and health conditions. We sought to assess whether certain types of MA plans serving subgroups of the MA population may be differentially impacted by FFS coding discrepancies, and therefore effected differently by CMS' proposal to eliminate the FFS Adjuster; depending on their enrollee mix, some MA plans could experience a higher or lower underpayment than the 7.74% mean underpayment that Avalere found in its March study. The differential impact based on enrollee

3 HHS-OIG, *Some Diagnosis Codes that Essence Healthcare, Inc., Submitted to CMS Did Not Comply With Federal Requirements* (Apr. 2019).

4 Avalere's use of CMS' methodology for purposes of testing alternative assumptions does not indicate agreement with the methods used in the CMS/CPI Study.

5 Predicted scores from each of the Avalere calibrated models correlate highly (>0.95) with scores from the CMS HCC payment model.

6 Avalere did not apply CMS' IPARS adjustment in its prior study, because the agency did not disclose that adjustment in its original and first supplemented release. Avalere did not apply it in this study, because it essentially eliminates the very difference in FFS and MA documentation standards that the CMS study purports to determine. We are unaware of any use of this adjustment in the literature or in CMS' past analyses. Avalere may offer additional comment on the fundamental problems and biases introduced by this adjustment at a later date.

mix could be increased by the alternative RADV approaches that CMS adopted for the CON14 audit (i.e., audits focused on specific beneficiary subcohorts).

In this study, Avalere estimated audit miscalibration estimates for subcohorts of the MA program:

- **Dual-Eligible Beneficiaries**: Avalere evaluated the impact of FFS coding discrepancies on MA plan payment for plans based on the number of beneficiaries eligible for both Medicare and Medicaid enrolled in the plan. Accepting CMS' methodology except for the adjustment to use actual numbers of claims, Avalere estimates the mean underpayment would be approximately 9.1% for a plan with 100% duals, compared with a mean underpayment of 7.3% for a plan with no duals.
- **Beneficiaries with Certain Common and Significant Health Conditions (e.g., diabetes):** Avalere examined beneficiaries with a condition in 10 disease severity hierarchies and 3 conditions that are not in a hierarchy (13 comparison groups) and compared beneficiaries in the comparison group to those who did not have that disease. For 2 of the diseases included in the study, the mean underpayment could be lower than overall average (e.g., the mean underpayment for cancer is estimated to be 6.6%). However, for several hierarchical condition categories (HCCs), it was substantially higher: 13.5% for kidney disease, 11.5% for CHF, and 10.3% for diabetes.

Overall, Avalere's analysis finds that there is significant variation in the audit miscalibration estimates based on the types of enrollees a plan serves. These findings have important implications given the statutory mandate that MA payments must be actuarially equivalent to the FFS program.[7]

Background

Payments to MA plans must be adjusted "so as to ensure actuarial equivalence" with the costs incurred in the traditional Medicare program.[8] Adjustments are required for several risk factors, including age, disability status, institutional status, and health status. To effectuate these requirements, CMS pays MA plans capitated amounts per member per month (PMPM) on a risk-adjusted basis using the CMS Hierarchical Condition Category (HCC) risk adjustment model. Under the model, risk scores are determined by demographic characteristics and diagnoses. MA plans substantiate their enrollees' health conditions by submitting diagnostic data to CMS.

In prior years, CMS audited a sample of MA plans annually through RADV, generating national- and contract-level estimates of payment error. Because the CMS-HCC risk adjustment model is primarily driven by enrollees' health conditions, RADV focuses on diagnostic coding errors (e.g., a diagnosis code that is not supported by a beneficiary's medical records).

7 42 U.S.C. § 1396w-23(a)(1)(C)(i).
8 Ibid.

CMS' HCC model is calibrated using FFS claims, which do not undergo the same auditing process as data reported by MA plans. To account for model mis-estimation as a result of underlying errors in the FFS claims data, CMS announced in a 2012 policy statement that it would apply a FFS Adjuster when calculating payment recoupment in RADV audits. This FFS Adjuster was intended to account for the different documentation standards used in creating the MA risk adjustment model, on the one hand, and in RADV audits, on the other. In practice, the FFS Adjuster sets a permissible level of payment error due to unsupported diagnosis codes and limits RADV audit recovery to payment errors above that level.[9]

In the 2012 RADV methodology announcement, CMS noted that it would conduct an audit of medical records accompanying FFS claims to determine the amount of the FFS adjuster. Although the amount of the FFS Adjuster had not been released by CMS, an internal CMS study had estimated that diagnosis coding errors in the FFS claims data had a negative impact on MA payment risk scores of between 4.8% and 8.1%.

2018 CMS Study of the Impact of the FFS Coding Discrepancies

In connection with the November 2018 Proposed Rule, CMS reviewed 8,630 FFS medical records,[10] adjusted the model calibration data based on an error rate derived from those records, and then re-estimated its HCC model. CMS refers to the difference in risk scores produced by the re-estimated model and its baseline as the "audit miscalibration estimate." The CMS study reports that, when applied to the MA population, the risk scores from the re-estimated model were almost equal to those produced by the original model (the re-estimated risk scores are 0.08% lower indicating slight overpayment to the plans with the original model). Therefore, CMS concluded erroneous coding in FFS had minimal impact on MA risk scores and no adjustment to payment recoupment is warranted under the RADV methodology (i.e., CMS proposed to eliminate the FFS Adjuster from its RADV methodology).

In its study, CMS made several simplifying assumptions that restricted variation in the data and underestimated the payment impact of FFS coding discrepancies. In particular, CMS adjusted the *claim-level HCC error rate* assuming: 1) every beneficiary with a particular HCC has the average number of claims supporting that HCC; and 2) each claim supporting an HCC has a probability of error equal to the average probability of error for that HCC overall. Using these assumptions, CMS computed *beneficiary-/HCC-level error rates* for each HCC. The methodology employed by CMS generated beneficiary-level error rates considerably lower than the claim-level error rates (see Tables 2a and 2b in the technical appendix of CMS' study), with the error rates for 15 of the 70 HCCs in that model being less than one-tenth of 1% (whereas claim-level error rates all exceed 20%). CMS used these adjusted beneficiary-level error rates in its effort to identify the audit miscalibration estimate.

9 A federal district court decision from September 2018 found that this type of adjustment was required to meet the statutory requirement of actuarial equivalence (UnitedHealthcare Ins. Co. v. Azar, 330 F. Supp. 3d 173 (D.D.C. 2018)..That decision is subject to appeal.

10 CMS has not fully disclosed why it selected this sample of FFS records for review or how the sample was created. Accordingly, Avalere has used the CMS reported HCC error rates as the basis of our study but is not able to comment on how the selection of these records may have biased CMS' analysis.

March 2019 Avalere Analysis Based on Alternative Beneficiary Claim Count Assumption

In a March 2019 study, Avalere tested alternatives to CMS' model. Avalere first evaluated the effect of using each beneficiary's actual number of claims per HCC, as reflected in the 5% sample for 2010[11] rather than assigning the average number of claims per HCC to all beneficiaries. Because this approach captures variation in the data by incorporating the relationship between the total number of claims and beneficiary costs, it should more closely estimate the real-world payment impact of coding and documentation errors in the FFS program.

Using the actual number of claims per beneficiary per HCC, rather than the average,[12] Avalere's study indicates that the impact of the FFS coding discrepancies on risk scores is considerably larger than the estimate in CMS' study—a 7.74% underestimate compared to CMS' estimate of -0.08% (i.e., a slight overestimate) (See Table 1). Following CMS convention from the CPI Study, negative percent differences reflect overpayment and positive percent difference reflect underpayment. Specifically, Avalere's analysis indicates that a root cause of the difference in the estimates between Avalere and CMS of the audit miscalibration is CMS' use of the average number of claims in calculating HCC person-level error rates, which significantly underestimates the error rate. In fact, most beneficiaries have 3 or fewer claims triggering the HCC; some beneficiaries have 1 or 2 claims triggering the HCC (see Appendix Table 1).

Table 1. Comparison of Audit Miscalibration Estimates, CMS Study and Alternative Beneficiary Claim Count Assumption

	Mean Underpayment* (Percent Difference)	95% Lower Bound	95% Upper Bound
CMS Study	-0.08%	-0.09%	-0.07%
Alternative Beneficiary Claim Count Assumption	7.74%	7.42%**	8.00%

*The CMS study findings are as reported in the CMS analysis. Therefore, a negative amount in Table 4 reflects an overpayment while a positive amount reflects an underpayment.
** In the March 2019 Study, Avalere inadvertently reversed the 95% Lower and Upper Bounds which are corrected here.

11 2010/2011 were the calibration years that CMS used for the V22 model

12 Avalere retained CMS' second key assumption for purposes of testing this alternative. Thus, similar to CMS, Avalere assumed, solely for the sake of analysis, that an error for an HCC on 1 claim for a beneficiary was independent of whether there was an error on another claim for the same HCC for that beneficiary.

Impact of Audit Miscalibration by Subpopulation

The impact of the audit miscalibration error found in our March 2019 study may vary across Medicare eligibility and disease group subpopulations. Payments to MA plans with vulnerable populations that do not reflect actuarial equivalence, as required by statute, may have negative effects for those plans and those beneficiaries. FFS coding discrepancies could ultimately result in larger mean underpayments to plans than the 7.74% Avalere found in its original analysis under certain circumstances.

In this follow-on study, Avalere conducted sensitivity analyses to determine how selected patient characteristics might affect estimates of the audit miscalibration error for populations that differ from the Medicare FFS average (e.g., those who are sicker or lower income). Avalere examined the impact of recalibration under its alternative assumption (that an individual's claim count should determine the probability of error) on certain subpopulation cohorts: the dual-eligible population and certain disease cohorts. Specifically, Avalere recomputed the audit miscalibration estimate for each group and compared it to the estimate for those not included in the group (e.g., duals versus non-duals, lung disease versus no lung disease). Avalere further decomposed the results of the hierarchy level analysis and calculated audit miscalibration estimates for the least severe and most severe disease within each hierarchy.

Analysis of Miscalibration Error by Dual-Eligibility Subgroup

Within the Medicare population, there are different subgroups based on Medicare eligibility status: aged (i.e., beneficiaries who are 65 or older), disabled, and dual-eligible (i.e., beneficiaries who are aged or disabled and who also qualify for Medicaid based on their income or health status). In this analysis, Avalere estimated the impact of audit miscalibration for the dual-eligible population compared to the non-dual-eligible population (see Table 2).[13] Avalere found that the dual-eligible enrollee subpopulation had a significantly higher audit miscalibration error than the non-dual subpopulation (over 1.7 percentage points), demonstrating that MA plans with more dual-eligible enrollees (e.g., dual-eligible special needs plans (D-SNPs) would experience larger underpayments using this model than plans with fewer dual-eligible enrollees. A plan that exclusively enrolls dual-eligible beneficiaries, such as a D-SNP would be expected to experience a 24% greater payment impact than a plan with no dual-eligible enrollees.

Table 2. Mean Underpayments Related to Audit Miscalibration Error; Dual-Eligibility Subgroups

	Non-Dual-Eligible Beneficiaries	Dual-Eligible Beneficiaries
Medicaid Status	7.34% (n=1,156,145)	9.08% (n=286,610)

Note: Values are statistically significant at 95% level (all p values <.0001)

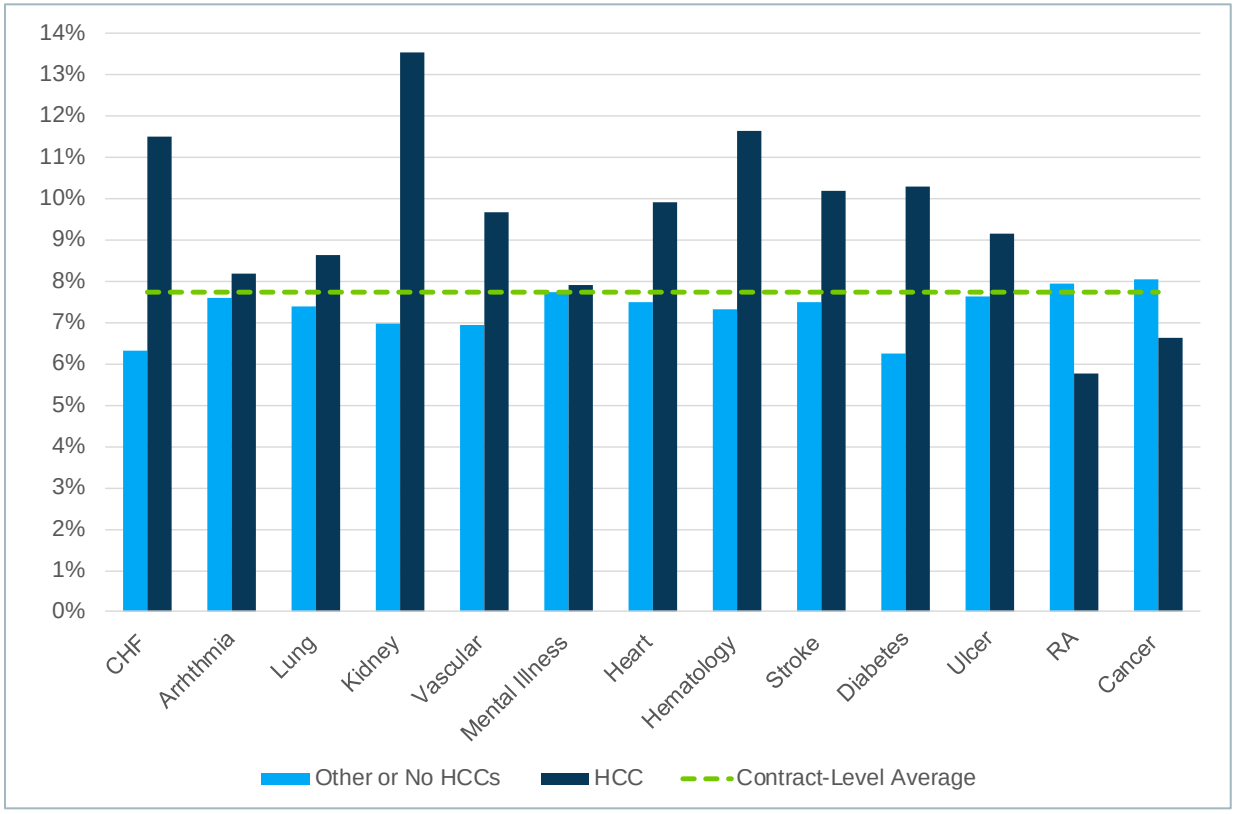
13 Note: The historical data that CMS, and subsequently Avalere, used for the analysis do not include an indicator for dual-eligibility status. Avalere therefore uses Low-Income Subsidy (LIS) status as a proxy for dual-eligibility

Analysis of Miscalibration Error by Select Disease Groups

Avalere also estimated the range of potential payment impacts for common disease groups by estimating the miscalibration error for certain HCCs. Specifically, Avalere estimated the mean underpayment related to audit miscalibration error for the entire hierarchy (e.g., all beneficiaries with an HCC for diabetes), for beneficiaries with the least severe version of a condition (e.g., HCC 19, diabetes without complication), and for beneficiaries with the most severe version of a condition (e.g., HCC 17, diabetes with acute complications). For its analysis of each HCC, Avalere also estimated the mean underpayment related to audit miscalibration error for all beneficiaries who did not have that HCC (e.g., beneficiaries who had either no HCCs or HCCs other than those within the diabetes hierarchy). Avalere also included 3 HCCs in the analysis that are not in a severity hierarchy (CHF, Arrythmia, and Rheumatoid Arthritis (RA)).

Avalere found substantial variation based on HCC and also based on condition severity (see Figure 1). For example, mean underpayment for HCCs ranged from 5.7% for RA to 13.5% for kidney disease. Within diabetes, mean underpayment for the least severe HCC was 6.7%, compared to 26.2% for beneficiaries with the most severe HCC (see Table 3).

Figure 1. Mean Underpayment Related to Audit Miscalibration Error as a Percent Difference



Only 2 diseases of the 13 evaluated had an audit miscalibration error below the population average of 7.7%, and both had miscalibration error above 5%. These findings demonstrate that, using this model, plans could experience significantly different payment outcomes based on the disease status of their enrollees, particularly under CMS' new methodology for conducting more targeted RADV audits.

Table 3. Mean Underpayments Related to Audit Miscalibration Error by Select Disease Subgroups

HCC*	Other or No HCCs	Entire HCC Group	Least Severe HCC	Most Severe HCC
Diabetes (HCCs 17, 18, 19)	6.2%	10.3%	6.7%	26.2%
Vascular (HCCs 107, 108)	6.9%	9.7%	9.4%	10.9%
Lung (HCCs 111, 112)	7.4%	8.6%	11.0%	8.4%
Cancer (HCCs 8, 9, 10, 11, 12)	8.0%	6.6%	5.9%	4.5%
Mental Illness (HCCs 57, 58)	7.7%	7.9%	8.3%	6.6%
Heart (HCCs 86, 87, 88)	7.5%	9.9%	8.9%	12.5%
Kidney (HCCs 135, 136, 137)	7.0%	13.5%	14.6%	11.9%
Stroke (HCCs 99, 100)	7.5%	10.2%	10.3%	9.4%
Hematology (HCCs 46, 48)	7.3%	11.6%	12.0%	10.5%
Ulcer (HCCs 157, 158, 161)	7.6%	9.1%	10.0%	5.4%
Arrhythmia (HCC 96)	7.6%	8.2%	n/a	n/a
CHF (HCC 85)	6.3%	11.5%	n/a	n/a
RA (HCC 40)	7.9%	5.7%	n/a	n/a

*The HCC hierarchies are listed according to the prevalence of the condition (the least severe HHC within each hierarchy is the highest number) followed by the 3 single HCCs, also listed according to prevalence. Values are statistically significant at 95% level (all p values <.0001)

Conclusion

Avalere's March 2019 analysis indicated that errors in the FFS claims used to calibrate CMS' HCC model have an impact on the accuracy of the model and, ultimately, on payment to plans. These results imply that the audit miscalibration may significantly impact the payment recoupment estimates generated by either CMS' traditional contract level RADV methodology or the revised methodology focusing on individuals and subcohorts. In this follow-on study, Avalere found that the effect would be significantly greater for the dual-eligible population and for certain HCCs. This means that plans with higher concentrations of dual-eligible beneficiaries or with beneficiaries with certain conditions would experience a higher rate of underpayment and highlights methodological concerns with CMS' study, which does not consider any variations in the impact of FFS coding discrepancies across different MA plans.

Methodology

For this analysis, Avalere used the Medicare 5% Limited Data Set (LDS) Standard Analytic Files (SAFs) for 2010 and 2011, the years used for the original calibration of the 2014 CMS-HCC model version 22 (V22). These data consist of a 5% longitudinal sample of Medicare beneficiaries that includes most information from all medical claims incurred by beneficiaries covered by FFS Medicare Parts A or B during these years. This database also includes demographic and Medicare enrollment information for these beneficiaries. Avalere also used summary information provided by CMS on CMS-HCC version 12 (V12) claim diagnosis error rates in its FFS Adjuster technical appendix to conduct the simulation analyses described below.

Avalere used the clinically revised model in use for payment between 2014 and 2016 (V22) for this analysis. Avalere mapped the HCC error rates reported in CMS' technical appendix Table 2a using information from the CMS 2011 paper on the CMS-HCC clinical revision to map the HCCs from V12 to V22.[14] CMS used model V12 in their analysis and mapped the error rates used as the basis for the audit miscalibration to V12 condition categories (CCs). CMS has since created a major clinical revision of the HCCs used in the model and has used 3 versions of the clinically-revised CMS-HCC models for payment between 2013 and 2018 (V21, V22, V23). All of these models have been used in payment to health plans at some point between 2011 and 2018.

Avalere used these data to estimate the range of potential effects on CMS-HCC model weights and risk scores under alternative assumptions for simulating the impact of diagnosis errors on the MA risk adjustment model. Avalere first identified Medicare beneficiaries in the 5% SAF sample who satisfied the inclusion criteria for calibrating the 2014 CMS-HCC V22 community model leg (i.e., beneficiaries continuously enrolled in Parts A and B, with no months of MA

14 Pope et al. "Evaluation of the CMS-HCC Risk Adjustment Model." CMS ORDI, March 2011. https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/downloads/Evaluation_Risk_Adj_Model_2011.pdf.

coverage, for 12 months and beneficiaries continuously enrolled in Parts A and B, with no months of MA coverage, for any number of months but including January 2015).

Avalere then identified the months in 2011 during which the beneficiary was not eligible for end-stage renal disease (ESRD) benefits, did not have institutionalized status, and did not incur a claim for which Medicare was the secondary payer. Indication of coverage for the Part D Low-Income Subsidy (LIS) was used as a proxy for dual eligible status due to limitations of the LDS SAFs. An indicator for whether a beneficiary incurred a nursing facility/custodial care facility evaluation and management (E&M) service claim submitted by a physician or non-physician practitioner for the current month and at least the three immediately previous months was used as a proxy for institutionalized status.

Next, Avalere calculated the sum of Medicare program payments (excluding inpatient indirect medical education capital and operating payments) for each qualifying beneficiary-month in 2011. Avalere identified ICD-9-CM diagnosis codes from qualifying claims for these beneficiaries incurred during 2010 for 1) inpatient and facility outpatient setting claims (using type of bill criteria); and 2) for physician and non-physician professional claims.

Avalere used the CMS-HCC V22 ICD-9-CM diagnosis code algorithms to determine the associated V22 Condition Category (CC). For each alternative error simulation model, Avalere repeated the following steps 250 times:

- With replacement, randomly sampled the community model-qualifying beneficiaries
- In this sample, randomly (based on the algorithms for each simulation alternative) selected person-claim-CCs as simulated errors and excluded from subsequent steps
- Created person-level indicators of HCC assignment (using CMS-HCC hierarchy logic) and other model indicators
- Estimated the 2014 V22 CMS-HCC community model using the sample generated as described above, applying all non-negativity and HCC hierarchy monotonicity constraints on estimated weights, using relative (to sample average) annualized payments
- Retained the estimated model weights and summary of the distribution of risk scores for each model

As a test of face validity, Avalere examined the impact of recalibration on our sample, the Avalere generated model predicted risk scores correlated at 0.95 or greater with the CMS-HCC model suggesting a high degree of fit between the models.

To generate the audit miscalibration estimates in the current study Avalere compared risk scores for subgroups of Medicare enrollees using coefficients from the CMS-HCC v22 model and Avalere's error simulation model. Avalere ran CMS' risk adjustment software to indicate the array of risk factors for each individual MA enrollee, producing dichotomous variables for each of the demographic and clinical factors. The sample included 1.4 million records, accounting for over 16.2 million member-months.

Next, Avalere estimated aggregate risk scores using the different sets of coefficients for subpopulations based on Medicaid status and disease group. For each subpopulation, Avalere restricted the sample to just patients of interest, then calculated the sum of the CMS-HCC v22 dichotomous variables weighted by months enrolled. The 2 sets of coefficients were then applied to the resulting table of risk factors to estimate risk scores. Avalere estimated the difference between the resulting risk scores in percentage terms to determine the audit miscalibration error.

Avalere assessed the following subgroups:

- Medicaid Status (i.e., dual-eligible and non-dual-eligible)
- Disease Hierarchies
    - Cancer (HCCs 8, 9, 10, 11, and 12), Diabetes (HCCs 17, 18, and 19), Heart (HCCs 86, 87, and 88), Hematology (HCCs 46 and 48), Kidney (HCCs 135, 136, and 137), Lung (HCCs 111 and 112), Mental Illness (HCCs 57 and 58), Stroke (HCCs 99 and 100), Ulcer (HCCs 157, 158, and 161), Vascular (HCCs 107 and 108) (See Appendix Table 2 for included diseases)
- Individual HCCs
    - Arrythmia (HCC 96), CHF (HCC 85), and Rheumatoid Arthritis (HCC 40)

Avalere compared the audit miscalibration error for subgroups of beneficiaries with and without each status, hierarchy, or HCC of interest. In the case of the disease hierarchies, Avalere additionally estimated the audit miscalibration error for patients with the most severe (i.e., lowest numbered) and least severe (i.e., highest numbered) HCCs.

Appendix

Appendix Table 1. Distribution of Claims per HCC and Beneficiary-Level Error Estimation, Top 10 CCs[15]

Version 22 CC	Adjusted CMS Claim Error Rate[16]	Avalere Sample Beneficiaries w/ CC	Claims Count				Beneficiary Error Rate Estimated Using Claim Count Mean and Percentile			
			Mean	Percentile			Mean (CMS Method)	Percentile		
				25th	50th	75th		25th	50th	75th
CC 19 – Diabetes w/o Complication	20.2%	362,090	7.0	3	5	9	0.001%	0.827%	0.034%	0.000%
CC 108 – Vascular Disease[17]	33.1%	234,285	3.8	1	2	5	1.430%	33.135%	10.979%	0.399%
CC 111 – Chronic Obstructive Pulmonary Disease[18]	20.7%	206,580	5.4	1	3	6	0.019%	20.665%	0.883%	0.008%
CC 96 – Specified Heart Arrhythmias[19]	41.7%	192,371	9.1	2	5	13	0.035%	17.389%	1.261%	0.001%
CC85 – Congestive Heart Failure[20]	30.4%	181,053	6.2	1	3	7	0.065%	30.409%	2.812%	0.024%

15 For illustrative purpose for this table, for each HCC, Avalere estimated beneficiary-level error rates by raising the claim-level error rate to an exponent equal to the number of claims that triggered the HCC. CMS' approach used the average rather than the actual number of claims. Appendix Table 2 provides the distribution of claims per beneficiary to show how different claims counts yield different beneficiary errors and that CMS' approach likely underestimates the beneficiary errors per HCC.

16 The claim-level error rates are adjusted using a multilevel model to account for small sample sizes for several HCCs. Specifically, the model assumes that the error rate for a claim is composed of a rate in common across HCCs within a clinically-related HCC category plus an error rate specific to that HCC. The resulting empirical Bayes estimates of these error rates yield error rates very similar to CMS' reported claim-level error rates when the sample is large and rates similar to the overall average error rate for all HCCs and the average among the HCCs in the same category when the sample is small

17 HCC 105 in CMS' analysis

18 HCC 108 in CMS' analysis

19 HCC 92 in CMS' analysis

20 HCC 80 in CMS' analysis

Version 22 CC	Adjusted CMS Claim Error Rate[16]	Avalere Sample Beneficiaries w/ CC	Claims Count				Beneficiary Error Rate Estimated Using Claim Count Mean and Percentile			
			Mean	Percentile			Mean (CMS Method)	Percentile		
				25th	50th	75th		25th	50th	75th
CC 18 – Diabetes w/ Chronic Complications	18.1%	139,453	4.4	1	3	5	0.050%	18.053%	0.588%	0.019%
CC 12 – Breast, Prostate, and Other Cancers and Tumors[21]	43.6%	114,906	6.8	1	3	7	0.349%	43.559%	8.265%	0.298%
CC 58 – Major Depressive, Bipolar, and Paranoid Disorders[22]	51.3%	86,593	7.0	2	4	8	0.955%	26.300%	6.917%	0.478%
CC 40 – Rheumatoid Arthritis and Inflammatory Connective Tissues Disease[23]	26.9%	75,782	5.4	1	3	7	0.083%	26.940%	1.955%	0.010%
CC 100 – Ischemic or Unspecified Stroke[24]	57.0%	59,820	4.5	1	2	5	8.124%	57.039%	32.535%	6.038%

*Avalere mapped CMS-HCC v12 errors to v22 and divided by a 1.066 adjustment factor to get average CMS weights equal to average relative monthly payments so that all models have the same

21 HCC 10 in CMS' analysis
22 HCC 55 in CMS' analysis
23 HCC 38 in CMS' analysis
24 HCC 96 in CMS' analysis

Appendix Table 2. Sample Sizes Disease Subgroups

HCC (Ordered from Most to Lease Severe)	Other or No HCCs	Least Severe HCC	Entire HCC Group	Most Severe HCC
Diabetes (HCCs 17, 18, 19)	1,060,849	237,345	381,906	4,777
Vascular (HCCs 107, 108)	1,198,206	212,475	244,549	32,074
Lung (HCCs 111, 112)	1,209,920	21,181	232,835	211,654
Cancer (HCCs 8, 9, 10, 11, 12)	1,264,582	92,393	178,173	15,951
Mental Illness (HCCs 57, 58)	1,335,420	79,802	107,335	27,533
Heart (HCCs 86, 87, 88)	1,359,376	37,348	83,379	15,422
Kidney (HCCs 135, 136, 137)	1,376,959	8,821	65,796	48,444
Stroke (HCCs 99, 100)	1,377,140	58,240	65,615	7,375
Hematology (HCCs 46, 48)	1,381,563	51,973	61,192	9,219
Ulcer (HCCs 157, 158, 161)	1,400,045	37,871	42,710	1,889
Arrhythmia (HCC 96)	1,247,342	n/a	195,413	n/a
CHF (HCC 85)	1,257,604	n/a	185,151	n/a
RA (HCC 40)	1,365,481	n/a	77,274	n/a

Appendix Table 3: Adjusted HCC Error Rates Used in Avalere Simulation and HCC Description

Group	HCC	Description	Adjusted Error Rate
Diabetes	19	Diabetes without Complication	20.2%
	18	Diabetes with Chronic Complications	18.1%
	17	Diabetes with Acute Complications	18.1%
Vascular	108	Vascular Disease	33.1%
	107	Vascular Disease with Complications	33.3%
Lung	112	Fibrosis of Lung and Other Chronic Lung Disorders	33.4%
	111	Chronic Obstructive Pulmonary Disease	20.7%
Cancer	12	Breast, Prostate, and Other Cancers and Tumors	43.6%
	11	Colorectal, Bladder, and Other Cancers	43.6%
	10	Lymphoma and Other Cancers	38.1%
	9	Lung and Other Severe Cancers	39.2%
	8	Metastatic Cancer and Acute Leukemia	31.0%
Mental Illness	58	Major Depressive, Bipolar, and Paranoid Disorders	51.3%
	57	Schizophrenia	43.6%
Kidney	137	Chronic Kidney Disease, Severe (Stage 4)	36.2%
	136	Chronic Kidney Disease, Stage 5	27.1%

	135	Acute Renal Failure	36.2%
Stroke	100	Ischemic or Unspecified Stroke	57.0%
	99	Cerebral Hemorrhage	43.0%
Hematology	48	Coagulation Defects and Other Specified Hematological Disorders	48.8%
	46	Severe Hematological Disorders	48.8%
Heart	88	Angina Pectoris	45.0%
	87	Unstable Angina and Other Acute Ischemic Heart Disease	33.5%
	86	Acute Myocardial Infarction	32.2%
Ulcer	161	Chronic Ulcer of Skin, Except Pressure	21.0%
	158	Pressure Ulcer of Skin with Full Thickness Skin Loss	25.0%
	157	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	25.0%
CHF	85	Congestive Heart Failure	30.4%
Arrhythmia	96	Specified Heart Arrhythmias	41.7%
RA	40	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	26.9%

Source: Medicare 5% Limited Data Set (LDS) Standard Analytic Files (SAFs) for 2011; CMS-HCC model version 22

Avalere mapped v12 error rates in CMS technical appendix to v22 HCCs

About Us

Avalere is a vibrant community of innovative thinkers dedicated to solving the challenges of the healthcare system. We deliver a comprehensive perspective, compelling substance, and creative solutions to help you make better business decisions. As an Inovalon company, we prize insights and strategies driven by robust data to achieve meaningful results. For more information, please contact info@avalere.com. You can also visit us at **avalere**.com.

Contact Us

Avalere Health
An Inovalon Company
1350 Connecticut Ave, NW
Washington, DC 20036
202.207.1300 | Fax 202.467.4455
avalere.com