



Medicare Payment
Advisory Commission

425 I Street, NW • Suite 701
Washington, DC 20001
202-220-3700 • www.medpac.gov

Michael E. Chernew, Ph.D., Chair
Amol Navathe, M.D., Ph.D., Vice Chair
Paul B. Masi, M.P.P., Executive Director

March 1, 2024

Chiquita Brooks-LaSure
Administrator
Centers for Medicare & Medicaid Services
Department of Health and Human Services
P.O. Box 8016
Baltimore, MD 21244-8016

Attention: CMS-2024-0006

Dear Ms. Brooks-LaSure:

The Medicare Payment Advisory Commission (MedPAC) welcomes the opportunity to comment on the Centers for Medicare & Medicaid Services (CMS) notice of proposed rulemaking entitled “Advance Notice of Methodological Changes for Calendar Year (CY) 2025 for Medicare Advantage (MA) Capitation Rates and Part C and Part D Payment Policies,” published on January 31, 2024. We appreciate your staff’s work on the notice, particularly considering the competing demands on the agency.

Our comments focus on the following provisions:

- CMS hierarchical condition category (CMS-HCC) risk-adjustment model for calendar year 2025
- MA coding pattern difference adjustment
- Normalization factors for CMS-HCC risk-adjustment models
- Normalization factors for CMS’s prescription drug hierarchical condition category risk-adjustment model

CMS hierarchical condition category (CMS-HCC) risk-adjustment model for calendar year 2025

Medicare payments to MA plans are adjusted to account for differences in enrollees’ expected medical spending. The purpose of risk adjustment is to ensure that plans are adequately and fairly compensated for treating all categories of enrollees—those with high expected medical spending as well as those likely to incur lower spending. The CMS-hierarchical condition category (CMS-HCC) risk-adjustment model uses demographic information (e.g., age, sex, Medicaid enrollment, and disability status) and certain diagnoses

to calculate a risk score for each enrollee. Diagnoses associated with similar medical conditions and with similar spending are grouped into HCCs. Some closely related HCCs are grouped into hierarchies based on condition severity.

Each demographic and HCC component in the risk-adjustment model has a coefficient that represents the expected medical expenditures associated with that component. These coefficients are estimated using fee-for-service (FFS) Medicare claims data such that all FFS Medicare spending in a year is distributed among the model components. The sum of these dollar value coefficients is normalized into an index, called a risk score. Normalization, applied to all risk scores for FFS and MA enrollees, establishes a 1.0 risk score for a beneficiary with expected spending equal to average FFS Medicare spending. Higher risk scores generate higher payments because beneficiaries with higher risk scores are expected to have higher expenditures; similarly, lower risk scores generate lower payments.

For 2024, CMS began phasing in a new risk-adjustment model version (V28) that is based on more recent FFS data and uses a new International Classification of Diseases, 10th Revision (ICD-10) diagnosis code-to-HCC mapping and an updated list of HCCs. (The prior model version was V24.) Revising the model with more recent data and a new diagnosis-to-HCC mapping will reflect more recent FFS diagnostic, FFS utilization, and FFS spending patterns.

The V28 risk-adjustment model also excludes or constrains the coefficients of certain HCCs that CMS identified as being subject to intentional or unintentional discretionary coding variation or inappropriate coding by health plans or providers. To do this, CMS identified the conditions in the model where coding in MA is highest relative to FFS and reviewed those conditions with clinical experts to determine if the coding differential likely indicates coding variation due to differences in diagnostic discretion. As a result of this process, CMS removed HCCs for protein-calorie malnutrition, angina pectoris, and atherosclerosis of arteries of the extremities with intermittent claudication. At the same time, CMS eliminated the levels of severity for two 3-HCC hierarchies, effectively combining each set of three HCCs into a single group with the same coefficient. CMS applied the hierarchy constraint to diabetes (with glycemic, unspecified, or no complications; with chronic complications; and with severe acute complications) and congestive heart failure (heart failure, except end-stage and acute; acute heart failure (excludes acute on chronic); and acute on chronic heart failure).¹

CMS began to phase in the new V28 model in 2024 by using a 33 percent / 67 percent risk score blend of the V28 / V24 models. For 2025, CMS proposes to continue phasing in the V28 model by using a 67 percent / 33 percent risk score blend of the V28 / V24 models.

Comment

MedPAC supports CMS's proposal to continue to phase in the V28 risk-adjustment model for 2025 payment to MA plans. Basing the model on more recent years of FFS data and using

¹ We interpret this proposal as an indication that CMS's analysis found that coding of the higher-severity HCCs in these hierarchies (rather than coding of any HCC in each hierarchy) was greater in MA relative to FFS. We also note that the three congestive heart failure HCCs are the lower three severity levels of a 6-HCC hierarchy that also includes heart failure with heart assist device/artificial heart; end-stage heart failure; and heart transplant status/complications.

ICD-10 diagnosis codes to estimate model coefficients are important updates that will improve the accuracy of risk adjustment and payments to MA plans.

We particularly support CMS's efforts to reduce the effects of coding differences between MA plans and FFS Medicare providers through the V28 risk model. Last year, in our comment on CMS's proposal to implement the V28 model for 2024, we reported higher rates of discretionary coding for eliminated or constrained HCCs (relative to FFS coding) among MA contracts with higher overall coding intensity. As we noted last year, when diagnostic discretion, intentional or unintentional, leads to large differences in the coding rates between MA and FFS and across MA plans, it diminishes the accuracy of risk-adjusted payments to MA plans and increases the payments that MA plans receive due to higher coding intensity. Eliminating HCCs and constraining the coefficients of HCCs that are found to have excessive discretionary coding variation is a reasonable strategy to improve payment accuracy and reduce overall MA coding intensity relative to FFS. We encourage CMS to continue efforts to identify other HCCs with discretionary or inappropriate MA coding and eliminate or constrain those HCCs, giving appropriate consideration to any effects on the accuracy of the risk-adjustment model overall and for beneficiaries with discretionary HCCs.

MA coding pattern difference adjustment

As noted above, Medicare payments to MA plans are adjusted to account for differences in enrollees' expected medical spending using the CMS-HCC risk-adjustment model. The model uses FFS Medicare claims data to estimate the model coefficients. Therefore, the model calculates an expected spending amount based on Medicare spending and diagnostic coding patterns for FFS beneficiaries. Most diagnoses for FFS beneficiaries are reported on physician and hospital outpatient claims. In FFS Medicare, such claims tend to be paid based on procedure codes, so there is little financial incentive for providers to document diagnoses that are not the primary reason for the visit. If certain diagnoses are not reported on FFS claims, the cost of treating those conditions is attributed to other components in the model, including the coefficients for age and sex categories.

Because FFS spending and diagnoses are used to calibrate the CMS-HCC risk model, payments to MA plans assume similar levels of diagnostic coding between FFS and MA. But MA plans have significant financial incentives to code more diagnoses (and have the tools to do so), because greater coding intensity can result in higher HCCs for their enrollees, which leads to higher payments from Medicare.

Since 2010, the Congress has required CMS to adjust MA enrollees' risk scores to reduce the impact of MA and FFS Medicare coding differences. An adjustment reducing MA risk scores by 3.41 percent was applied from 2010 through 2013. Starting in 2014, legislation specified a minimum reduction of about 4.9 percent, which rose gradually to a minimum adjustment of about 5.9 percent in 2018, where it will remain until the Secretary implements a risk-adjustment model that uses MA diagnostic, MA utilization, and MA plans' cost data.² To date,

² Section 1853 (a)(1)(C)(ii) of the Social Security Act [42 U.S.C. 1395w-23(a)(1)(C)(ii)].

the Secretary has opted to reduce MA risk scores by the minimum amount required by law. For 2025, CMS proposes once again to apply the minimum required adjustment of 5.9 percent.

Comment

We understand the agency's inclination to proceed cautiously in making large changes to Medicare payments—whether provider payments under FFS, or payments to health plans under Medicare Advantage—in order to ensure that beneficiary access is not compromised. However, for nearly a decade, the Commission has documented increased payments to MA plans due to coding intensity and has recommended policies to address the problem. We estimate that, in 2024 alone, coding intensity will increase Medicare's payments to MA organizations by about \$50 billion dollars. The cost of those increased payments is borne by taxpayers and by the Medicare beneficiaries and state Medicaid agencies who help finance the Medicare program through Part B premiums. We estimate that Part B premiums will be higher for all Medicare enrollees by about \$8 billion in 2024 because of that increased spending.

The evidence documented by the Commission and others over many years indicates that stronger action is needed. Although we applaud CMS's efforts to reduce MA and FFS coding differences by eliminating or constraining certain HCCs in the risk-adjustment model proposed for 2024, those efforts are inadequate to address growing MA coding intensity and the resulting increased payments to MA plans. The Commission's recommendation (outlined below) and other proposals to more fully address coding intensity are within the Secretary's authority. Given the fiscal burden on the taxpayers and beneficiaries who finance the Medicare program, it is imperative that CMS act now to better account for the effects of coding intensity.

CMS's adjustment does not fully account for coding differences

Over the past several years, a growing body of research has demonstrated that the impact of MA and FFS coding differences are far larger than the minimum adjustment that the Secretary has routinely applied. At least eight independent studies, using a variety of methods and data sources, corroborate the Commission's estimates that the impact of plan

coding intensity has always been larger than the adjustment that CMS applied in any given year.^{3,4,5,6,7,8,9,10}

In the advance notice of payment rates for 2019, CMS requested comment on adopting an alternative method for calculating the MA coding adjustment factor, including the Commission's existing cohort method and the demographic estimate of coding intensity (DECI) method.¹¹ Kronick and Chua developed the DECI method and used it to produce estimates of coding intensity that are double the estimates previously produced by the Commission. For example, Kronick and Chua estimated that coding intensity in 2019 was 20.0 percent, whereas we previously estimated it at 10.0 percent.^{4,12} We analyzed both methods to better understand the reasons for the differing coding intensity estimates. We found that, for 2019, the different estimates were largely due to differences in the data used and the treatment of Medicaid eligibility and new MA enrollees.¹³

- after revising our original cohort method to account for differences in Medicaid eligibility and remove a restriction on beneficiaries being continuously enrolled in the same program (MA or FFS), we estimated MA coding intensity to be 12.4 percent (up from 10.0 percent when using our original cohort method), and
- after revising the DECI method to use complete beneficiary-level data, account for differences in Medicaid eligibility, and constrain new enrollees, we estimated MA coding intensity to be 11.6 percent (down from 20.0 percent as published by Kronick and Chua).

³ Kronick, R., and F. M. Chua, Department of Health and Human Services. 2021. *Industry-wide and sponsor-specific estimates of Medicare Advantage coding intensity*. November 11. Available at <https://ssrn.com/abstract=3959446>.

⁴ Jacobs P. D., and R. Kronick. 2018. Getting what we pay for: How do risk-based payments to Medicare advantage plans compare with alternative measures of beneficiary health risk? *Health Services Research* 53(6): 4997-5015.

⁵ Hayford, T. B., and A. L. Burns. 2018. Medicare Advantage enrollment and beneficiary risk scores: Difference-in-differences analyses show increases for all enrollees on account of market-wide changes. *Inquiry* 55 (January-December): 46958018788640.

⁶ Congressional Budget Office. 2017. *Effects of Medicare Advantage enrollment on beneficiary risk scores*. Working paper 2017-08. Washington, DC: CBO.

⁷ Geruso, M., and T. Layton. 2020. Upcoding: Evidence from Medicare on squishy risk adjustment. *Journal of Political Economy* 12, no. 3 (March): 984-1026.

⁸ Kronick, R., and W.P. Welch. 2014. Measuring coding intensity in the Medicare Advantage program. *Medicare & Medicaid Research Review* 4, no. 2.

⁹ Government Accountability Office. 2013. *Medicare Advantage: Substantial excess payments underscore need for CMS to improve accuracy of risk score adjustments*. GAO-13-206. Washington, DC: GAO.

¹⁰ Government Accountability Office. 2012. *Medicare Advantage: CMS Should improve the accuracy of risk score adjustments for diagnostic coding practices*. GAO-12-51. Washington, DC: GAO.

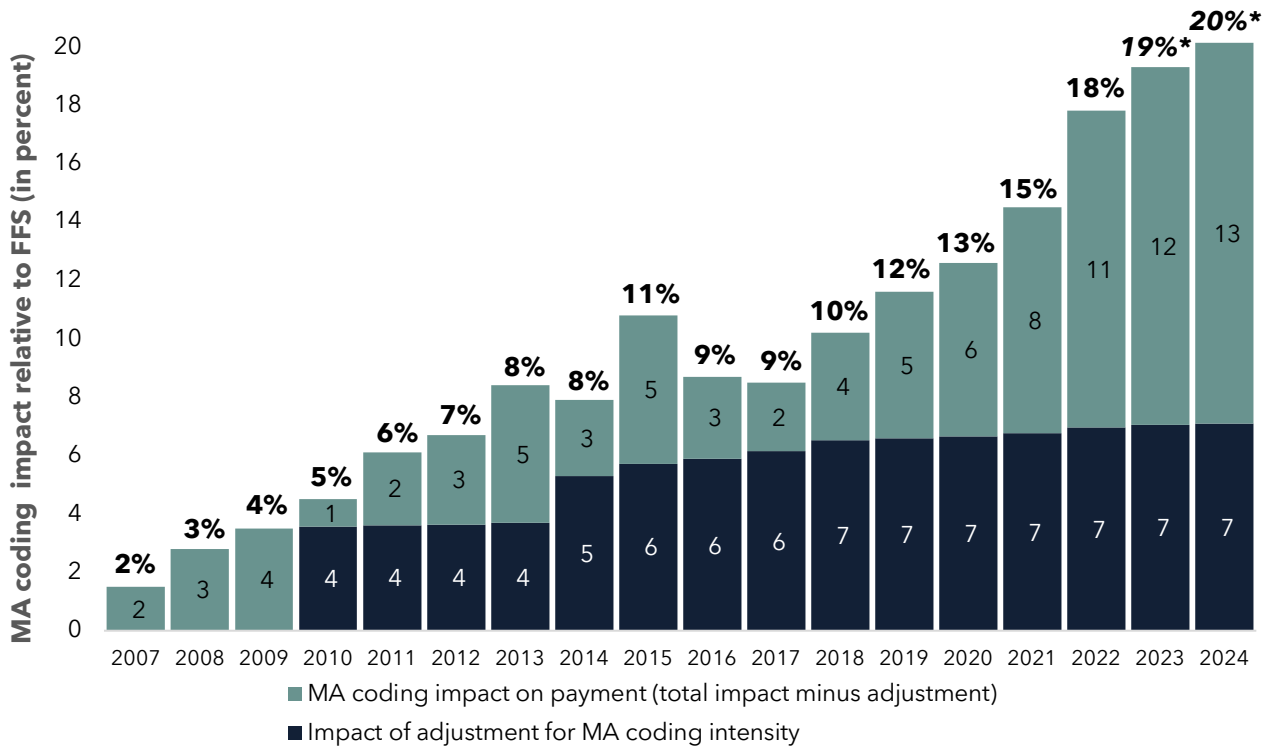
¹¹ Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2018. Advance notice of methodological changes for calendar year (CY) 2019 for Medicare Advantage capitation rates, Part C and Part D payment policies and 2019 draft call letter. February 1. <https://www.cms.gov/medicare/health-plans/medicareadvtspeccratestats/downloads/advance2018.pdf>.

¹² The Commission's original cohort method estimates presented in this letter are reported as a percent of the average FFS risk score and therefore are different from the original cohort method estimates we previously published, which were reported as a percent of the average MA risk score. For example, we previously reported that the share of MA risk scores attributable to coding intensity in 2019 was 9.1 percent, which is equivalent to saying that MA risk scores are 10.0 percent higher than FFS risk scores for comparable beneficiaries ($1 / (1 - 0.091) = 1.100$ or 10.0 percent). Changing the denominator from average MA risk score to average FFS risk score allows for direct comparison with the DECI method and does not reflect a change in the magnitude of our estimates.

¹³ This analysis is described in more detail in Chapter 13 of our forthcoming March 2024 report to the Congress.

Thus, we found that addressing these issues in the cohort and DECI methods resulted in estimates of MA coding intensity that were very similar (within 1.5 percentage points) for all years 2008 through 2022. The consistency of the results using two different methods of estimation bolsters confidence in the accuracy of the estimates. Because the revised DECI method allows us to include a larger share of MA and FFS beneficiaries (including those with institutional status and partial years of enrollment), we now report estimates of MA coding intensity based on this method rather than the cohort method. In Figure 1, the percentages at the top of each bar show our estimate of the impact of coding intensity for each year. The lower, dark blue portion of each column shows the effect of the coding adjustment that CMS applied in each year and the upper, teal portion shows the level of uncorrected coding intensity resulting in higher payments to MA plans.¹⁴

Figure 1 | Estimated impact of coding intensity on MA risk scores was larger than coding adjustment, 2007-2024



Note: MA (Medicare Advantage), FFS (fee-for-service). All estimates account for any differences in age, sex, Medicaid eligibility, and institutional status between MA and FFS populations. New enrollees are constrained to have no coding intensity as their risk scores are not based on diagnostic coding. The annual adjustment for MA coding began in 2010. MA coding

¹⁴ Note that the impact of the annual coding adjustment has been updated since our January 2024 public presentation. The statutory coding intensity adjustment is denominated as an adjustment to MA risk scores. Because the DECI method produces estimates of coding intensity that are expressed as a percentage difference in FFS risk scores, we change the coding intensity adjustment so that it can be applied appropriately. We calculate the impact of the 5.9 percent reduction to MA risk scores for 2024 as $1.20 \times 5.9\% = 0.071$ or 7.1%. As a result, for each year in which the coding adjustment is applied, the uncorrected coding intensity and the associated payment amounts to MA plans are lower than we reported in our January 2024 public presentation.

intensity has increased MA risk scores annually, but increases were offset by new versions of the risk-adjustment model in 2014, 2016, and 2017 and by increased FFS coding in 2016 and 2017. The impact of the coding adjustment is calculated as the MA coding intensity estimate relative to FFS, multiplied by the coding adjustment. For 2024, we calculate $1.20 \times 5.9\% = 0.071$ or 7.1%.

*For 2023 and 2024, we project coding intensity based on the annual trend from 2017 through 2021, an increase of 1.5 percentage points per year. For 2024, we reduced the annual trend by 0.67 percentage points to account for one-third of an estimated 2 percentage point reduction in coding intensity associated with the introduction of the V28 risk-adjustment model, which will be phased in over three years.

Source: MedPAC analysis of CMS enrollment and risk score files.

Our estimates of coding intensity in 2023 and 2024 are projected based on the annual trend from 2017 through 2021, when coding intensity increased by an average of 1.5 percentage points per year. For 2024, we reduced the annual trend by 0.67 percentage points to account for one-third of an estimated 2 percentage point reduction in coding intensity associated with the introduction of a new risk-adjustment model (V28), which will be phased in over three years.¹⁵ However, there is uncertainty about the impact of moving to the V28 on MA coding intensity. We will continue to monitor those effects and update our analysis as needed.

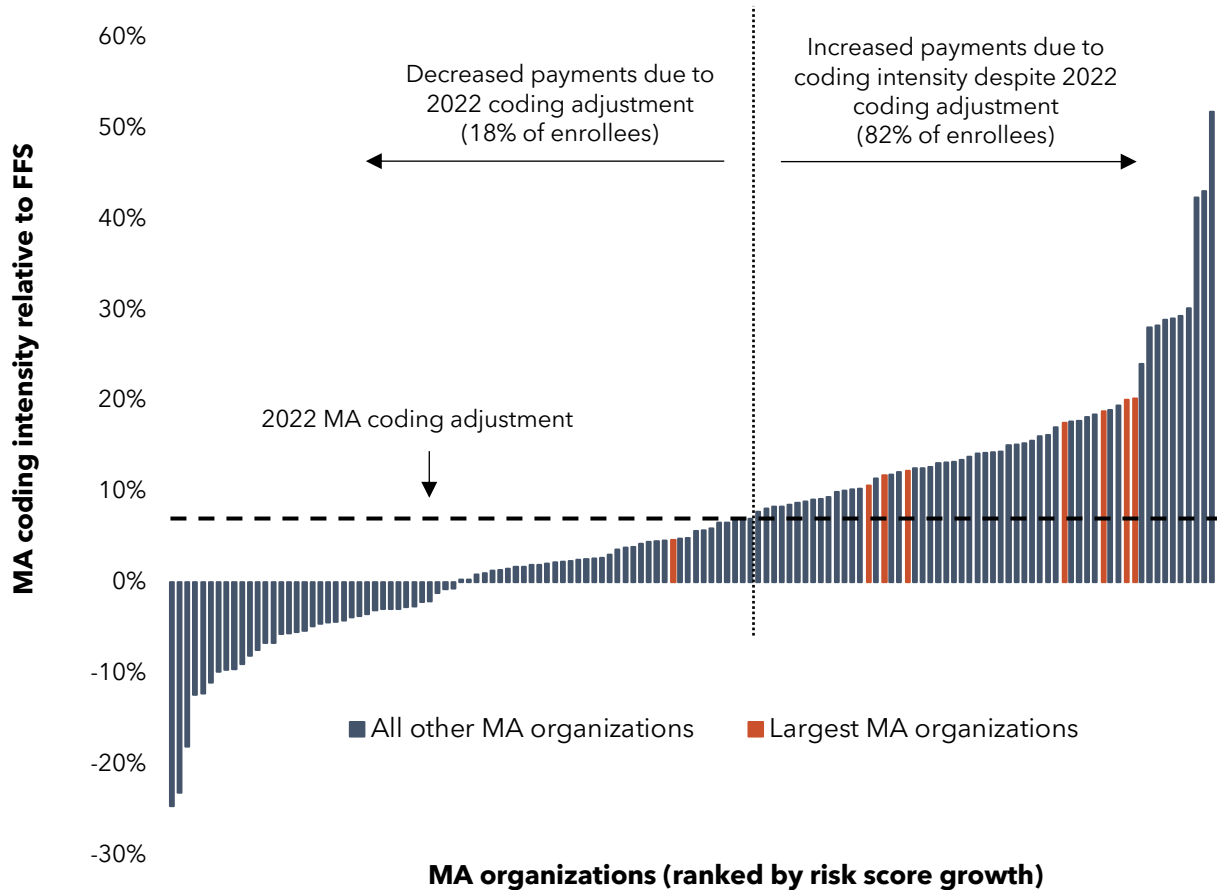
For 2024, we project that MA risk scores will be about 20 percent above risk scores for comparable FFS beneficiaries, far more than the impact of CMS's coding intensity adjustment. As a result, we estimate that unaccounted-for coding intensity will increase MA risk scores relative to FFS risk scores by 13 percent in 2024, resulting in about \$50 billion in increased payments to MA plans.

Applying an across-the-board coding-intensity adjustment generates payment inequity across MA contracts and organizations

We estimated the effect of coding intensity separately for each MA organization and found substantial variation in coding for 2022. Figure 2 shows coding intensity relative to FFS coding by MA parent organization, excluding contracts in the Program of All-Inclusive Care for the Elderly, special needs plans, and organizations with fewer than 2,500 enrollees.

¹⁵ The new risk-adjustment model (V28) introduced in 2024 is expected to reduce MA risk scores relative to FFS because it removes HCCs or constrains the coefficients of HCCs that have much higher rates of MA coding relative to FFS. We believe a 2 percent reduction in risk scores under the V28 model relative to the previous V24 model is a reasonable expectation based on a combination of factors. In 2014, CMS implemented a model that similarly removed HCCs or constrained HCC coefficients with higher MA coding rates, which reduced MA risk scores by roughly 2.0 percentage points or 2.5 percentage points, depending on the year. Also, in the 2024 advance notice, CMS reported that the combined effect of changing from the V24 to the V28 model and of the normalization factor for 2024 would reduce payments to MA plans by 3.12 percentage points. We note that the average annual effect of the normalization factor over the past five years is -2.1 percentage points, with somewhat smaller effects in more recent years. There is uncertainty about the impact of moving to the V28 on MA coding intensity. We will continue to monitor those effects and will update our analysis as we are able.

Figure 2 | Coding intensity varied across MA organizations relative to FFS, 2022



Note: MA (Medicare Advantage), FFS (fee-for-service). Excludes special needs plans, contracts for the Program of All-Inclusive Care for the Elderly, and organizations with fewer than 2,500 enrollees in the analysis. All estimates account for any differences in age, sex, Medicaid eligibility, and institutional status between MA and FFS populations. New enrollees are constrained to have no coding intensity because their risk scores are not based on diagnostic coding.

Source: MedPAC analysis of CMS enrollment and risk score files.

Consistent with prior years, we found that coding intensity varied significantly across MA organizations in 2022. We estimate that about half of MA organizations (covering 18 percent of MA enrollees) had coding intensity that was lower than CMS’s 2022 coding adjustment; by contrast, the other half of organizations (covering 82 percent of MA enrollees) had coding intensity that increased their payments from Medicare even after the 2022 coding adjustment was applied. These differences demonstrate that CMS’s across-the-board adjustment for coding intensity, which reduces all risk scores by the same amount, generates inequity across contracts by reducing net revenue for plans with lower coding intensity and allowing other plans to retain a significant amount of revenue from higher coding intensity.

We also found significant variation in coding intensity across the largest eight MA organizations (covering 77 percent of MA enrollees), ranging from about 5 percent to about 20 percent above FFS levels. Seven of the eight largest MA organizations had greater coding intensity than the 2022 coding adjustment and therefore received increased payments due to their coding practices. These differences are large enough to give MA organizations with higher coding intensity a significant competitive advantage by increasing the size of plan rebates and helping them to attract more enrollees.

MA plans have several tools that are unavailable in FFS to identify diagnoses

MA plans use several mechanisms that do not exist in FFS Medicare to document diagnoses for their enrollees. They can identify enrollees likely to have an HCC that has not yet been documented using data the plan already has: an enrollee's historical claims, risk score data, and prescription drug data (e.g., a prescription for insulin likely indicates a diabetes diagnosis). Of all the mechanisms to document more diagnosis codes, evidence continues to highlight MA plans' use of health risk assessments and chart reviews as major sources of plan revenue from coding intensity.

MA plans' use of health risk assessments to identify additional diagnoses

Health risk assessments are provided to Medicare beneficiaries as part of an annual wellness visit, and, for MA enrollees, health risk assessments are often provided during a plan-initiated home visit.¹⁶ Health risk assessments sometimes rely on patients self-reporting medical conditions, which may result in HCCs based on inaccurate diagnoses, diagnoses that are no longer active (and therefore not eligible for risk adjustment), or diagnoses without sufficient evidence to conform to ICD coding guidelines.¹⁷ (More information about these concerns is in our March 2023 report chapter on Medicare Advantage.¹⁸)

We analyzed 2021 encounter records to identify HCCs that were supported only by a health risk assessment, meaning that there was no physician or hospital service provided to treat a beneficiary for a specific health condition during the same calendar year.¹⁹ In 2021, about 6.9 million MA enrollees had a health risk assessment that identified at least one HCC, and a total of 15.0 million unique HCCs were identified through health risk assessments. Of those, 3.2 million beneficiaries had a health risk assessment that was the only source for at least one of the HCCs identified, and a total of 5.0 million HCCs (one-third of all HCCs identified

¹⁶ Recent reporting shows that agents and brokers are often paid by plans to conduct health risk assessments of new enrollees, but such assessments are not allowable for risk adjustment because agents and brokers are not clinicians.

¹⁷ Department of Justice. 2022. *United States files civil fraud lawsuit against Cigna for artificially inflating its Medicare Advantage payments*. Washington, DC: DOJ. <https://www.justice.gov/usaosdny/pr/united-states-files-civil-fraud-lawsuit-against-cignaartificially-inflating-its>.

¹⁸ See Chapter 11, Medicare Payment Advisory Commission. 2023. Report to the Congress: Medicare payment policy. Washington, DC: MedPAC.

¹⁹ The general steps we followed were to identify physician and hospital encounter records allowable for risk adjustment; identify each record as a health risk assessment (using procedure codes for annual wellness visit or initial preventive physical exam, or an evaluation and management visit provided in the home), chart review (using chart review indicator), or other service; map diagnoses from those records to HCCs; apply HCC hierarchies; compare the HCCs we identified from encounter records with the HCCs in CMS's risk score file and exclude HCCs not identified in both sources; apply HCC coefficients for the appropriate risk model; and apply Part A and Part B payment rates specific to each plan. We then identified the number of HCCs and associated dollar amounts that were supported through a health risk assessment, chart review, or both.

on health risk assessments) were identified only on a health risk assessment. Seven HCCs each generated more than \$500 million in payments from these assessments (based on the frequency with which they were documented on HRAs and the amount Medicare pays on average for each), accounting for nearly 60 percent of all payments generated by health risk assessments.²⁰ We found that in 2022, diagnostic coding that was associated with only health risk assessments accounted for \$13 billion in payments to MA plans, or a little more than 3 percent of all payments to MA plans (about \$4.7 billion, or roughly one percentage point of spending, reflects HCCs documented on both HRAs and on chart reviews; more information about this is in the note to Figure 3). About 60 percent of these payments were from health risk assessments conducted as part of an annual wellness visit or initial preventive physical examination, while the rest of these payments were from in-home health risk assessments.^{21,22}

MA plans' use of chart reviews to identify more diagnoses

Some MA plans devote significant effort to conducting chart reviews to increase MA payments.²³ Chart reviews allowable for risk adjustment document the diagnoses made during hospital and physician encounters in which medical services were provided. MA plans use chart reviews to identify diagnoses not captured through the usual means of reporting diagnoses (e.g., claims data and encounter data): diagnoses that are not reported on the provider's claim sent to the MA plan, diagnoses made during an encounter in which the MA plan does not submit a record of the encounter to CMS, or diagnoses made during an encounter in which the total number of diagnoses from that encounter exceeds the number of diagnosis fields on the encounter record. Because Medicare requires each HCC to be supported by diagnostic evidence in a patient's medical record (chart), chart reviews are one way for plans to identify diagnoses not captured through provider claims or on plan encounter data. However, because chart review programs are used exclusively in MA (there is no financial incentive to undertake chart reviews in FFS Medicare), all diagnoses newly documented through chart reviews contribute to differences in FFS and MA diagnostic coding and increase payments to MA plans.

Like health risk assessments, some MA plans treat chart review programs as an independent revenue stream that yields a positive return on investment because the

²⁰ The seven HCCs that each generated more than \$500 million in payments from health risk assessments and the percentage of the time that a health risk assessment was the only source of the HCC were vascular disease, 47 percent; major depressive, bipolar, and paranoid disorders, 46 percent; morbid obesity, 38 percent; chronic obstructive pulmonary disorder, 25 percent; diabetes with chronic complications, 15 percent; coagulation defects and other specified hematological disorders, 60 percent; and congestive heart failure, 23 percent. We note that diabetes with chronic complications and congestive heart failure are among the HCCs that have a constrained coefficient in the V28 risk model, meaning that differences in the level of severity (e.g., diabetes without complications, with chronic complications, or with acute complications) for these conditions are not reflected in the V28 risk-adjustment model coefficients, which may contribute to an expected reduction in overall MA and FFS coding differences.

²¹ Office of Inspector General, Department of Health and Human Services. 2020. *Billions in estimated Medicare Advantage payments from diagnoses reported only on health risk assessments raise concerns*. OEI-03-17-00471. Washington, DC: OIG.

²² The impact of health risk assessments on payments to MA plans has grown. For 2017, the Office of Inspector General (OIG) found that HCCs supported only by a health risk assessment accounted for \$2.6 billion, or about 1.3 percent of payments to MA plans in 2017. OIG excluded beneficiaries with more than one health risk assessment during the year. Our analysis included all health risk assessments. For 2020, we found that HCCs supported only by a health risk assessment accounted for \$8.6 billion, or about 2.9 percent of all payments to MA plans.

²³ Office of Inspector General, Department of Health and Human Services. 2021. *Some Medicare Advantage companies leveraged chart reviews and health risk assessments to disproportionately drive payments*. Washington, DC: OIG.

additional Medicare payments from newly documented diagnoses far exceed the costs of paying nurses and medical assistants to review medical charts.²⁴ Several lawsuits allege that MA plans use chart reviews to identify new diagnosis codes, but not to verify the accuracy of already submitted codes. Some lawsuits allege that an MA organization is aware that diagnoses submitted to CMS are not supported by the medical chart and therefore violated Medicare's rules governing the reporting of diagnoses.^{25,26,27} Some plans and vendors appear to selectively review charts with a higher likelihood of increasing revenue and use artificial intelligence to more accurately identify likely revenue-producing charts.^{28,29} While the financial return is worth plan sponsors' effort and financial investment, chart review programs increase the financial burden for the taxpayers and beneficiaries (in both MA and FFS) who fund the Medicare program.

We analyzed 2021 MA encounter records to identify HCCs that were supported by a chart review but not through any other record of a physician or hospital encounter during the same calendar year. In 2021, about 11.6 million (more than 40 percent of) MA enrollees had a chart review that identified at least one HCC, and a total of 32.7 million unique HCCs were identified on MA chart reviews. Of the 11.6 million MA enrollees with a chart review, about half (5.9 million) had a chart review that was the only source of an HCC, and a total of 9.6 million HCCs (about 30 percent of all HCCs identified on chart reviews) were identified only through a chart review. Eight HCCs each generated more than \$1 billion in Medicare payments from chart reviews (based on the frequency with which they were documented on chart reviews and the amount Medicare pays on average for each), accounting for more than half of all chart review-based payments.³⁰ We found that, in 2022, chart reviews alone increased Medicare's payments to MA plans by about \$25 billion, and accounted for about 7 percent of all payments to MA plans (about \$4.7 billion, or roughly 1 percentage point, reflects HCCs documented on both chart reviews and on health risk assessments; more information about this is in the note to Figure 3).³¹

²⁴ The legal complaints cited in this section support this statement. One complaint includes exhibits of plan documents that detail the financial performance of the plan's chart review program (United States of America v. Anthem 2020).

²⁵ United States of America ex rel. Benjamin Poehling v. UnitedHealth Group Inc., et al. 2016. U.S. District Court for the Central District of California. No. 11-cv-0258-A. <https://dockets.justia.com/docket/virginia/vaedce/2:2020cv00079/467869>.

²⁶ United States of America ex rel. James M. Swoben v. Secure Horizons, et al. 2017. U.S. District Court for the Central District of California. No. CV 09-5013 JFW (JEMx).

²⁷ United States of America v. Anthem, I. 2020. US District Court for the Southern District of New York. No. 1:20-CV-02593. <https://www.justice.gov/usao-sdny/press-release/file/1262841/download>.

²⁸ Blue Health Intelligence. 2020. Predictive analytics to optimize coding accuracy. <https://www.bluehealthintelligence.com/riskadjustment/>

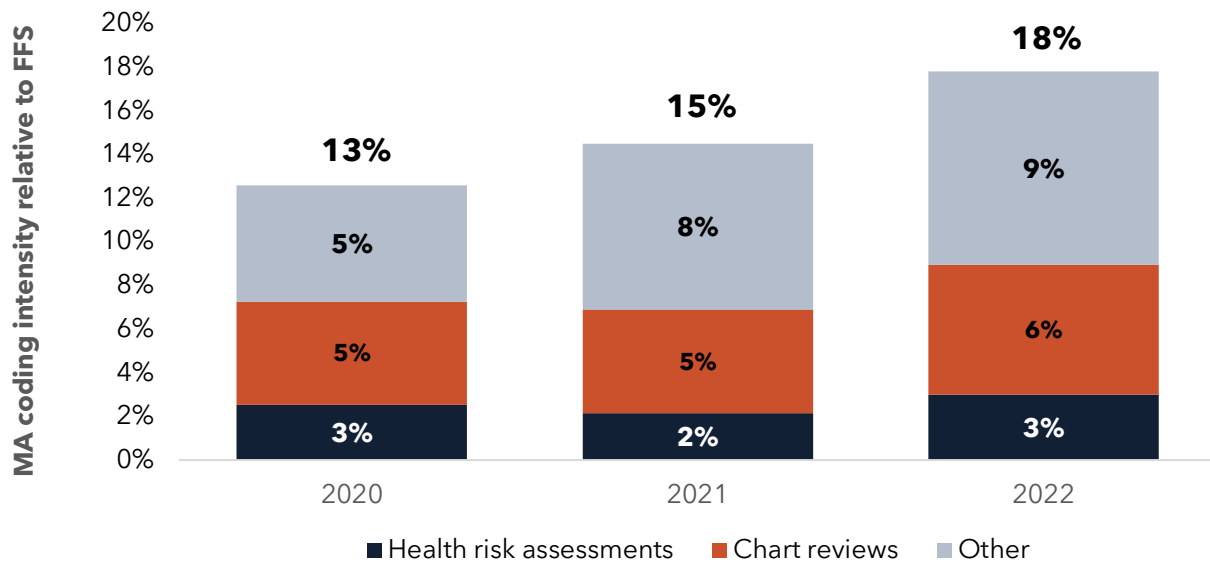
²⁹ Optum. 2020. *A smarter retrospective risk adjustment program*. Eden Prairie, MN: Optum.

³⁰ The eight HCCs that each generated more than \$1 billion in payments from chart reviews and the percentage of the time that a chart review was the only source of the HCC were vascular disease, 36 percent; chronic obstructive pulmonary disorder, 29 percent; major depressive, bipolar, and paranoid disorders, 36 percent; diabetes with chronic complications, 19 percent; congestive heart failure, 26 percent; morbid obesity, 37 percent; disorders of immunity, 45 percent; and rheumatoid arthritis and inflammatory connective tissue disease, 32 percent. We note that diabetes with chronic complications and congestive heart failure are among the HCCs that have a constrained coefficient in the V28 risk model, meaning that differences in the level of severity (e.g., diabetes without complications, with chronic complications, or with acute complications) for these conditions are not reflected in the V28 risk-adjustment model coefficients, which may contribute to an expected reduction in overall MA and FFS coding differences.

³¹ The impact of chart reviews on payments to MA plans has grown substantially in recent years. The Office of Inspector General found that HCCs supported only by a chart review accounted for \$6.7 billion in payments to MA plans, or about 3.2 percent of payments to MA in 2017. For 2020, we found that HCCs supported only by a chart review accounted for \$12.7 billion, or about 3.4 percent of all payments to MA plans. (Office of Inspector General, Department of Health and Human

We estimate that chart reviews and health risk assessments together accounted for about \$33 billion in payments to MA plans, or about 9 percent of all payments to MA plans in 2022.³² Combined with our finding that all sources of coding intensity resulted in MA risk scores that were about 18 percent higher than risk scores for comparable FFS beneficiaries in 2022, we conclude that health risk assessments and chart reviews together accounted for about 50 percent of all MA coding intensity (Figure 3). Both our analysis and that of the Office of Inspector General have found that the use of health risk assessments and chart reviews varies significantly across MA organizations.³³

Figure 3 | Chart reviews and health risk assessments accounted for about half of overall MA coding intensity, 2020-2022



Note: MA (Medicare Advantage), FFS (fee-for-service). Figure shows the impact of coding intensity on payments to MA plans for the years 2020 through 2022. The underlying diagnoses were reported during health care encounters in the prior year, 2019 through 2021, respectively. Spending associated with HCCs documented on both a chart review and a health risk assessment, but not during any other physician or hospital encounter during the same calendar year (about 1 percentage point of spending in 2022), have been distributed proportionately to the health risk assessment or chart review categories.

Source: MedPAC analysis of CMS enrollment and risk score files, and Medicare Trustees' reports.

Services. 2021. *Some Medicare Advantage companies leveraged chart reviews and health risk assessments to disproportionately drive payments.* Washington, DC: OIG.)

³² About \$4.7 billion in payments to MA plans—about 1 percentage point of spending—were from HCCs identified on both a health risk assessment and a chart review, but not during any record of a physician or hospital encounter during the same calendar year.

³³ Office of Inspector General, Department of Health and Human Services. 2021. *Some Medicare Advantage companies leveraged chart reviews and health risk assessments to disproportionately drive payments.* Washington, DC: OIG.

MedPAC's approach to addressing coding intensity

In our March 2016 report to the Congress, the Commission recommended a multipronged approach that would fully account for the impact of coding differences, improve the equity of the adjustment across MA contracts, and increase incentives to reduce costs and improve quality. The Commission's approach to reduce the impact of MA coding intensity has been to address the underlying causes first (e.g., remove health risk assessments and reduce year-to-year coding variations by using two years of diagnostic data) and then address remaining differences with either an across-the-board or tiered adjustment. The Commission's 2016 recommendation did not address the use of chart reviews because data were not available at that time, but eliminating chart reviews as a source of diagnoses for risk adjustment is consistent with the Commission's approach.

The recommendation, which would replace the existing mandatory minimum coding intensity adjustment (which has reduced MA risk scores by 5.9 percent since 2018), has three parts:

- develop a risk-adjustment model that uses two years of FFS and MA diagnostic data,
- exclude diagnoses that are documented only on health risk assessments from either FFS or MA, and then
- apply a coding adjustment that fully accounts for the remaining differences in coding between FFS Medicare and MA plans.

Implementing the first two policies—using two years of diagnostic data and excluding diagnoses documented through health risk assessments alone—and excluding chart review data from risk adjustment (consistent with the Commission's approach) would result in a more equitable, targeted adjustment to MA contracts than the current across-the-board adjustment. As noted earlier, health risk assessments and chart reviews alone account for roughly half of MA coding intensity. The Commission carefully considered options for addressing coding intensity and supports this approach because it balances implementation feasibility, administrative burden, and effectiveness.

Adjusting for any remaining coding intensity differences could also improve equity across MA contracts. Under one approach, contracts would be grouped into tiers of high, medium, and low coding intensity, and a coding intensity adjustment would be applied based on each tier's average level of coding intensity.³⁴ CMS has used a similar approach to select MA contracts for risk-adjustment data validation audits.³⁵ This policy would improve the overall equity of the coding intensity adjustment relative to the single, across-the-board adjustment used today. Finally, we note that in 2016, when the Commission voted on this recommendation, estimates of MA coding intensity net of CMS's coding adjustment were

³⁴ Medicare Payment Advisory Commission. 2016. Medicare Advantage: Calculating benchmarks and coding intensity. Presentation at the Commission's November public meeting. <http://medpac.gov/docs/default-source/meeting-materials/ma-ab-coding-nov16-for-laptop.pdf?sfvrsn=0>.

³⁵ For risk-adjustment data validation audits in 2011, CMS grouped all contracts into high, medium, and low levels of coding intensity and selected 20 high-level, 5 medium-level, and 5 low-level contracts at random.

much smaller than they are for 2024. Given that the impact of the Commission's recommendation, which would fully account for the effects of higher MA coding intensity, has grown substantially, policymakers could contemplate phasing in and monitoring the impact of the Commission's recommendation.

Normalization factors for CMS-HCC risk-adjustment models

The CMS-HCC risk-adjustment models are calibrated with diagnostic and spending information for beneficiaries enrolled in FFS Medicare. The risk-adjustment models are prospective in that they use health status in a base year (i.e., data collection year) to estimate incremental spending for a variety of beneficiary characteristics (e.g., age and gender) and health conditions in the following year (i.e., the payment year). To create relative factors, each model variable's incremental spending estimate, referred to as a dollar coefficient, is divided by the predicted average per capita expenditure for beneficiaries in FFS Medicare in a given year (i.e., the denominator year). Risk scores are the sum of relative factors assigned to each beneficiary based on their demographic characteristics and health status from the prior year. The average risk score is 1.0 among FFS beneficiaries in the denominator year.

The average FFS risk score changes each year due to an underlying trend that reflects changes in the health status and demographic characteristics of the population, and coding practices. Therefore, when a risk-adjustment model predicts expenditures in years other than the denominator year, the average FFS risk score may no longer be 1.0, as it was in the denominator year. Accordingly, an adjustment must be applied to account for the FFS risk score trend between the denominator year and payment year, and the adjustment (called "the normalization factor") must be specific to each model. For 2025, CMS proposes to use a blend of V24 and V28 models for CMS-HCC risk scores and thus needs to calculate normalization factors for each model.

In the past, CMS has used a 5-year linear trend of FFS risk scores to estimate a normalization factor for the payment year. However, CMS found that fluctuations in FFS risk scores during the COVID-19 pandemic required changes to the normalization factor estimation method. CMS found that the average FFS risk score (V28 model) changed -3.2 percent between 2020 and 2021, +2.5 percent between 2021 and 2022, and +1.7 percent between 2022 and 2023.

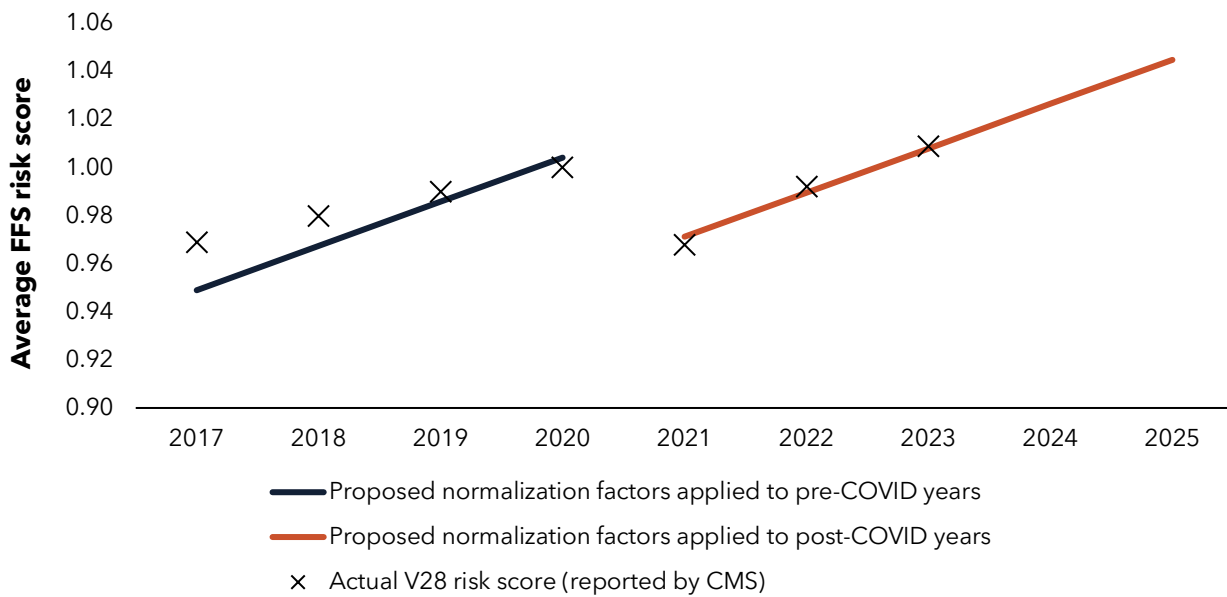
For 2024, CMS used data for 2018 through 2022, excluding 2021, to calculate the normalization factor. For 2025, CMS considered applying the same approach (using the five most recent years of data, excluding 2021) as well as other approaches (using the five most recent years of data, excluding 2021 and 2022; or using only post-COVID risk scores for 2021 to 2023), but all of these approaches yielded FFS risk score trends that were inconsistent with recent FFS risk score trends, and CMS concluded that those approaches were not supportable.

For 2025, CMS proposes using a regression approach using 2019 to 2023 risk scores with a COVID-19 flag applied to 2021 and later years. The approach uses risk scores for all five years, but accounts for the level change in FFS risk scores that is observed between 2020 and 2021. A separate regression analysis would be applied for the V24 and V28 models.

Comment

The Commission acknowledges the difficulty in predicting average FFS risk scores in light of the fluctuations that have occurred in recent years and supports CMS’s proposed regression approach for estimating the normalization factor for 2025. We find that the regression factors proposed for the V28 model fit actual FFS risk scores reasonably well, including for post-COVID years (Figure 4).

Figure 4 | CMS’s proposed normalization factor approach for the V28 model fits average FFS risk score data reasonably well, 2017–2025



Note: FFS (fee-for-service). COVID-19 (). Figure shows V28 model normalization factors (Intercept -36.1638, annual change in FFS risk score 0.0184, and COVID-19 flag -0.0513 as reported by CMS) applied to 2017 through 2025, and actual average FFS risk scores as reported by CMS for 2017 through 2023.

Source: MedPAC analysis of CMS normalization factor data for V28 model on page 65 of “Advance Notice of Methodological Changes for Calendar Year (CY) 2025 for Medicare Advantage (MA) Capitation Rates and Part C and Part D Payment Policies.”

Further, we note that CMS’s proposed regression approach is more cautious than relying on 2021 to 2023, post-COVID risk scores alone. If there are concerns that the normalization factors for 2025 will be too high when using the regression approach, leading to payments to MA plans being too low, we note some inaccuracy is inevitable under any prediction model, and it is likely that these inaccuracies will balance out over time. For example, CMS has reported that between 2015 and 2020, the normalization factor underestimated the actual

average FFS risk score by roughly 2 to 3 percentage points in each year, which translated directly into higher payments for MA plans.³⁶

Normalization factors for CMS’s prescription drug hierarchical condition category risk-adjustment model

In Part D, Medicare aims to subsidize about 75 percent of the cost of basic drug benefits, with enrollees paying the remainder (about 25 percent) through monthly premiums. Medicare’s subsidy takes the form of two distinct payments: 1) prospective capitated payments (the “direct subsidy”); and 2) cost-based reinsurance payments that cover a portion of an individual’s drug spending above the benefit’s out-of-pocket threshold (net of postsale rebates and discounts).³⁷

Direct subsidy payments are based on plan sponsors’ estimates of the expected costs of providing Part D’s basic benefit for an enrollee of average health. To limit plan incentives to engage in risk selection, direct subsidy payments are risk adjusted to reflect the expected costliness of each enrollee using the prescription drug hierarchical condition category risk-adjustment model (RxHCC model).

Similar to the CMS–HCC risk-adjustment model used to adjust payments to MA plans, the RxHCC model uses demographic information (e.g., age, sex, disability status) and certain diagnoses to adjust payments to Part D plans. Each demographic and RxHCC component in the risk-adjustment model has a coefficient that represents the expected drug costs associated with that component. These coefficients are estimated using the prescription drug event data for both stand-alone prescription drug plans (PDPs) and Medicare Advantage–Prescription Drug (MA–PD) plans, such that all gross Part D spending for basic drug benefits excluding reinsurance (“gross plan liability”) in a year is distributed among the model components. The sum of these dollar value coefficients is normalized into an index, called a risk score. Normalization accounts for underlying trends such as changes in demographic characteristics and diagnostic coding, and it establishes a 1.0 risk score for a beneficiary with average Part D spending across all Part D enrollees. Higher risk scores generate higher payments because beneficiaries with higher risk scores are expected to have higher drug spending, while beneficiaries with lower risk scores are expected to have lower drug spending.

The goal of the RxHCC model is to predict well for subgroups of Part D beneficiaries rather than for individual beneficiaries. This ensures that plans are paid accurately based on the average expected spending of their enrollees compared with a beneficiary with average

³⁶ See Table II-1 on page 44: Centers for Medicare & Medicaid Services, Department of Health and Human Services. 2020. Memo to all Medicare Advantage organizations, Prescription Drug Plan sponsors, and other interested parties regarding advance notice of methodological changes for calendar year (CY) 2022 for Medicare Advantage (MA) capitation rates and Part C and Part D payment policies–Part II.

³⁷ Beginning in 2024, annual premium increases are limited to no more than 6 percent; when this cap is binding and premiums are constrained more than they otherwise would have been without such a cap, the Medicare program’s subsidy will necessarily increase to above 75 percent of basic benefit costs. Beginning in 2030, in addition to the 6 percent cap on annual premium increases, the Medicare subsidy will be capped at 80 percent; when these two limitations are at odds with each other, the Medicare program subsidy cap will take priority.

Part D spending. CMS measures the predictive accuracy in the RxHCC model by taking the ratio of predicted cost to actual cost. A predictive ratio of 1.0 means that the model perfectly predicts plan spending on average for a group of beneficiaries. The agency generally considers a predictive ratio between 0.90 and 1.10 to be reasonably accurate.

Before 2016, the RxHCC model was calibrated based only on FFS diagnoses and Part D costs for PDPs. In 2016, CMS updated the RxHCC model such that it is now calibrated with cost and diagnostic data for beneficiaries enrolled in both MA-PD plans and PDPs. At the time, the agency noted that incorporating both FFS and MA-PD data “allows MA-PD plan coding and utilization patterns to be accurately reflected in the Part D relative costs and improves the predictive accuracy of the RxHCC model.” However, since then, CMS has found that “the increase in MA-PD enrollment combined with the different coding and cost patterns for enrollees in MA-PD plans and PDPs has resulted in a diverging trend in average MA-PD and PDP risk scores over time.” Specifically, the agency noted that the existing RxHCC model tended to *overpredict* MA-PD plan costs (a predictive ratio of 1.106) and *underpredict* PDP costs (a predictive ratio of 0.879). The agency also noted that “this differential puts upward pressure on standardized bids for PDPs and, as a result, creates an unlevel playing field that generally inhibits fair competition between MA-PD plans and PDPs.”

In 2025, CMS expects that the benefit redesign under the Inflation Reduction Act of 2022 will result in a significant increase in the Part D plan liability, the costs for which Part D plans will be at risk. Given the much greater importance of risk adjustment under the redesigned benefit structure, combined with the shift in Part D enrollment from PDPs to MA-PD plans, the agency proposes to apply separate normalization factors for MA-PD plans and PDPs. By using separate normalization factors, the agency expects risk scores to “more accurately reflect Part D costs in each of these two sectors of the Part D market that are driven by a variety of market-based variables, including the overall benefits that they are able to manage, the lack of an ability of PDPs to affect the submission of diagnoses in FFS, and available strategies used to manage Part D costs.”

Comment

MedPAC supports CMS’s proposal to apply separate normalization factors for MA-PD plans and PDPs to ensure that the risk scores estimated using the RxHCC model accurately reflect the relative spending across subgroups of beneficiaries. In particular, the Commission has noted that because “risk adjustment plays a vital role in a capitated payment system by counterbalancing plan incentives for selection and ensuring that plans receive adequate payment,” an accurate risk-adjustment model would be critically important under a redesigned benefit that increased plan liability.³⁸

For years, the Commission has been concerned about the coding difference between FFS and MA that contributes to inaccurate payment adjustment using the CMS-HCC risk-adjustment model (see the section on MA coding pattern difference adjustment). Given the

³⁸ Medicare Payment Advisory Commission. 2020. *Report to the Congress: Medicare and the health care delivery system*. Washington, DC: MedPAC. <https://www.medpac.gov/document/june-2020-report-to-the-congress-medicare-and-the-health-care-delivery/>.

overlap of conditions in the CMS-HCC and RxHCC models, higher coding intensity by MA plans for medical conditions could increase coding intensity for conditions that are also used in the RxHCC model. Unlike the CMS-HCC risk-adjustment model, the RxHCC model is calibrated using costs and diagnostic data for beneficiaries enrolled in both FFS and MA. However, when there are large, systematic differences between the two sectors in coding intensity and patterns of medication use (for a given diagnosis), the RxHCC model's accuracy for predicting costs for either sector would be compromised even if the model performed well on average across all Part D beneficiaries. The growing divergence in risk scores for MA-PD plans and PDPs seems to confirm the existence of a systematic difference in coding intensity. Further, differences in coding practices and utilization patterns between the two sectors could both contribute to lower predictive accuracies for PDPs and MA-PDs, and may have different policy implications.

As CMS noted, a risk-adjustment model that, on average, systematically underpays PDPs while overpaying MA-PDs can have a direct impact on the sustainability of the PDP market. This is because when the model underpredicts costs for PDPs, these plans face greater pressure to charge higher premiums than MA-PD plans whose costs are overpredicted, making PDPs less competitive. Using a separate normalization factor for the two sectors of the Part D market would help level the playing field between PDPs and MA-PDs. However, we encourage CMS to continue to monitor the trends in Part D risk scores to ensure that the systematic differences between the two sectors do not result in a financial disadvantage to PDPs that undermines Part D's market-based structure built on competition among private plans.

Conclusion

MedPAC appreciates your consideration of these issues. The Commission values the ongoing collaboration between CMS and MedPAC staff on Medicare policy, and we look forward to continuing this relationship. If you have any questions regarding our comments, please do not hesitate to contact Paul Masi, MedPAC's Executive Director, at 202-220-3700.

Sincerely,



Michael E. Chernew, Ph.D.
Chair

MC/aj/ss/toh