

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

Michael Chernew, Ph.D., Chair Paul B. Ginsburg, Ph.D., Vice Chair James E. Mathews, Ph.D., Executive Director

March 3, 2022

Chiquita Brooks-LaSure, Administrator Centers for Medicare & Medicaid Services Department of Health and Human Services 200 Independence Avenue, SW Washington DC, 20201

RE: CMS-2022-0021

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) 2023 for Medicare Advantage (MA) Capitation Rates and Part C and Part D Payment Policies," published on February 2, 2022. We appreciate your staff's work on the notice, particularly considering the competing demands on the agency.

Our comments focus on the provisions related to Medicare Advantage coding pattern adjustment.

Medicare payments to MA plans are adjusted to account for differences in enrollees' expected medical costs. The purpose of risk adjustment is to ensure that plans are adequately and fairly compensated for treating all categories of enrollees—those with high medical costs as well as those likely to incur fewer costs. 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 treatment costs are grouped into HCCs.<sup>1</sup>

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 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 average FFS Medicare spending. Higher risk scores generate higher payments because beneficiaries with high risk scores are expected to have higher expenditures and vice versa. Risk adjustment using diagnostic information

 $<sup>^{\</sup>rm I}$  Some closely-related HCCs are grouped into hierarchies based on condition severity.

creates a clear financial incentive for MA plans: Documenting more HCCs increases Medicare's payment to the MA plan.

Because the CMS–HCC model uses FFS Medicare claims data to estimate the model coefficients, the model calculates an expected spending amount based on FFS Medicare costs and diagnostic coding patterns. Most diagnoses are reported on physician and outpatient claims, which in FFS Medicare tend to be paid based on procedure codes, thus providing little financial incentive to document diagnoses for FFS beneficiaries. 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. For MA payments to be accurate, diagnoses must be coded with the same intensity in both FFS Medicare and MA. When MA plans submit more diagnoses for a beneficiary than would have been documented in FFS Medicare, the program spends more for that beneficiary in MA than it would have if the beneficiary were in FFS. Because MA plans have significant financial incentives to code as many diagnoses as possible, coding intensity is higher in MA than in FFS Medicare, and payments to MA plans are thus higher than intended.

The Deficit Reduction Act of 2005 (DRA) mandated that CMS study the impact of coding differences on MA payments and required the agency to make an adjustment to MA risk scores to address the impact of MA and FFS coding differences. The study concluded with an adjustment reducing MA risk scores by 3.41 percent, which 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 risk adjustment using MA diagnostic, cost, and use data.<sup>2</sup>

To date, the Secretary has reduced MA risk scores by the minimum amount required by law and has not produced another study of the impact of coding intensity. For 2023, 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 (positive or negative) to Medicare payments—whether provider payments under FFS or payments to health plans under Medicare Advantage—so as to ensure that beneficiary access is not compromised. These concerns may be heightened given the value that Medicare beneficiaries put on the MA program, evidenced by strong year-over-year growth in MA enrollment.

However, over time, coding intensity has generated tens of billions of dollars in excess payments to MA plan sponsors. The cost of those payments is borne by taxpayers, Medicare beneficiaries, and state Medicaid agencies who fund the Medicare program.<sup>3</sup> We assert that the evidence documented by MedPAC and others over many years indicates that stronger action to address coding intensity is needed. The current approach of making only the statutory minimum adjustment to MA plan payments is not consistent with current law, since the DRA states that in

<sup>&</sup>lt;sup>2</sup> Section 1853 (a)(1)(C)(ii) of the Social Security Act [42 U.S.C. 1395w–23(a)(1)(C)(ii)].

<sup>&</sup>lt;sup>3</sup> Many Medicare beneficiaries and state Medicaid agencies help fund the Medicare program through Part B premiums.

applying risk adjustment to payments for MA plans "the Secretary shall ensure that such adjustment reflects changes in treatment and coding practices in the fee-for-service sector and reflects differences in coding patterns between Medicare Advantage plans and providers under part A and B to the extent that the Secretary has identified such differences." Failure to stem the excess spending created by coding intensity further jeopardizes the Medicare program's already challenging fiscal sustainability. We urge the Secretary and CMS to increase the coding intensity adjustment to more fully reflect the magnitude of this excess spending.

The agency's current approach to addressing MA coding intensity is problematic for three reasons: (1) the adjustment does not account for the full impact of coding intensity and allows excess spending to continue, increasing the overall cost of the Medicare program, escalating Part B premiums for all beneficiaries, and accelerating the depletion of the Part A trust fund; (2) the across-the-board adjustment (reducing all MA risk scores by 5.9 percent) generates payment inequity between contracts coding more aggressively and those coding more similarly to FFS Medicare; and (3) the approach does not address the underlying causes of coding intensity, thereby allowing coding intensity to undermine the goal of plans competing on the basis of quality and costs.

# CMS's adjustment does not fully account for coding differences, inflating payments to MA plans by more than \$91 billion between 2007 and 2022

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 our findings that the impact of plan coding intensity has always been larger than the adjustment that CMS applied in any given year. 5,6,7,8,9,10,11,12

MedPAC has also been estimating the impact of coding intensity since 2015, and we are able to estimate the magnitude of MA coding intensity relative to Medicare's payment adjustment for

<sup>&</sup>lt;sup>4</sup> Section 1853 (a)(1)(C)(ii)(I) of the Social Security Act [42 U.S.C. 1395w–23(a)(1)(C)(ii)(I)].

<sup>&</sup>lt;sup>5</sup> 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 <a href="https://ssrn.com/abstract=3959446">https://ssrn.com/abstract=3959446</a>. <sup>6</sup> 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.

<sup>&</sup>lt;sup>7</sup> 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.

<sup>&</sup>lt;sup>8</sup> Congressional Budget Office. 2017. *Effects of Medicare Advantage enrollment on beneficiary risk scores*. Working paper 2017-08. Washington, DC: CBO.

<sup>&</sup>lt;sup>9</sup> Geruso, M., and T. Layton. 2015. *Upcoding: Evidence from Medicare on squishy risk adjustment*. NBER working paper no. 21222. Cambridge, MA: National Bureau of Economic Research.

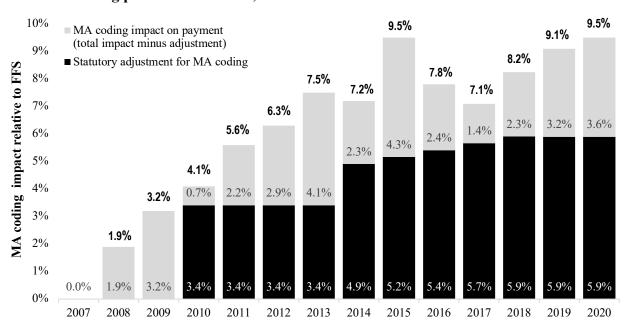
<sup>&</sup>lt;sup>10</sup> Kronick, R., and W. P. Welch. 2014. Measuring coding intensity in the Medicare Advantage program. *Medicare & Medicaid Research* Review 4, no. 2.

<sup>&</sup>lt;sup>11</sup> 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.

<sup>&</sup>lt;sup>12</sup> 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

each year from 2007 (the first year the CMS–HCC was fully implemented) to 2020 (the most recent year of available data). In Figure 1, the percentages at the top of each bar show our estimate of the impact of coding intensity for each year. The black portion of each column shows the coding adjustment that CMS applied in each year, while the gray portion shows the level of uncorrected coding intensity resulting in inflated payments to MA plans. Without intervention, MA coding intensity increases by about 1 percentage point per year, although the increase from 2019 to 2020 was smaller than in prior years. (Decreases in 2014, 2016, and 2017 were due to one-time events.<sup>13</sup>)

Figure 1. Impact of coding intensity on MA risk scores was larger than the adjustment for coding pattern differences, 2007–2020



Note: MA (Medicare Advantage), FFS (fee-for-service). All estimates account for any differences in age and sex between MA and FFS populations. Annual adjustment for MA coding began in 2010. MA coding intensity increased MA risk scores by about 1 percentage point annually, but was offset by new risk adjustment model versions in 2014, 2016, and 2017 and by increased FFS coding in 2016 and 2017.

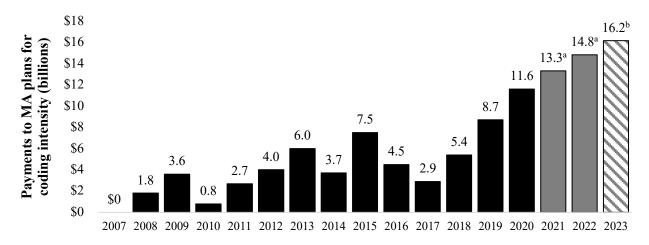
Source: MedPAC analysis of CMS enrollment and risk score files.

13

<sup>&</sup>lt;sup>13</sup> MA coding intensity fell in 2014, 2016, and 2017 due to the introduction of model versions that were less susceptible to MA and FFS diagnostic coding differences. In 2016 and 2017, MA risk scores grew at about the same rate as in prior years, but FFS risk scores grew at a faster rate, likely caused by Medicare's transition from using International Classification of Diseases (ICD)–9 to ICD–10 diagnosis codes in October 2015.

Exacerbating the effects of coding intensity—driven overpayments is the fact that beneficiary enrollment in MA is growing faster than ever. The combination of enrollment growth and coding intensity will result in excess Medicare spending of almost \$15 billion in 2022 alone (Figure 2). By the end of 2022, Medicare will have cumulatively paid MA plans more than \$91 billion just due to coding intensity since the CMS–HCC model was implemented. If CMS implements its proposal to apply the minimum 5.9 percent adjustment in 2023, we estimate that Medicare spending for coding intensity will rise by \$16.2 billion, to a total of more than \$107 billion since 2007. Given the financial status of the Medicare program, it is imperative that CMS act now to fully account for the impact of coding intensity.

Figure 2. Uncorrected MA coding intensity has generated more than \$91 billion in payments to plans since 2007 and will generate \$16 billion more under CMS's proposal for 2023



Note: MA (Medicare Advantage). Estimates for 2007 through 2020 are based on MedPAC's estimate of uncorrected coding intensity and Medicare spending for MA plans from the Medicare Trustees' Reports.

Source: MedPAC analysis of CMS enrollment and risk score files, and Medicare Trustee's Reports, 2019–2021.

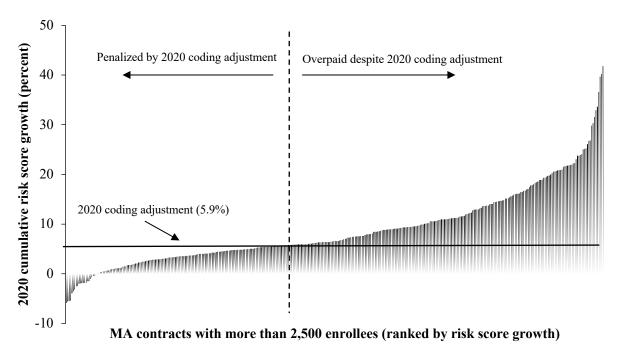
## Applying an across-the-board adjustment generates payment inequity across MA contracts

Each year, we break down our overall estimate of coding intensity impact relative to the statutory minimum coding adjustment and to FFS Medicare at the MA contract level (contracts may include one or more plans from the same sponsor).

<sup>&</sup>lt;sup>a</sup> The estimates for 2021 and 2022 conservatively assume that uncorrected coding intensity will be the same as in 2020 (3.6 percent, although all evidence suggests that it will be larger) and are based on projected Medicare spending for MA plans from the 2021 Medicare Trustees' Report.

<sup>&</sup>lt;sup>b</sup> The estimate for 2023 conservatively assumes that uncorrected coding intensity will be the same as in 2020 (3.6 percent, although all evidence suggests that it will be larger), is based on projected Medicare spending for MA plans from the 2021 Medicare Trustees' Report, and assumes that CMS will apply the proposed 5.9 percent coding adjustment.

Figure 3. Cumulative MA risk score growth relative to local FFS varied across contracts, 2020



Note: MA (Medicare Advantage), FFS (fee-for-service). Excludes MA contracts with enrollment below 2,500 (represents less than 1 percent of MA enrollment), contracts for the Program of All-Inclusive Care for the Elderly, and special needs plans. Analysis is based on retrospective cohorts of 2020 enrollees, tracked backward for as long as they were continuously enrolled in the same program (FFS or MA) or as far back as 2007.

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

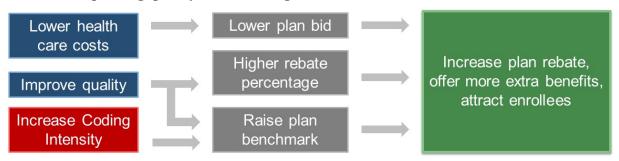
The substantial variation in coding intensity illustrated in Figure 3 highlights the second problem with CMS's approach to addressing coding intensity: The across-the-board adjustment generates inequity across plans, with plans that code diagnoses most aggressively continuing to financially benefit from that coding, while plans that code more similarly to FFS Medicare are penalized. These payment differences are substantial.

# CMS's approach does not address the underlying causes of coding intensity, allowing coding intensity to undermine the goal of plans competing on the basis of quality and costs

Incentives to improve quality and reduce health care costs are driven by the policy determining MA plan rebates. Rebates are one of the primary ways that MA plans compete because rebates fund the extra benefits (e.g., reduced cost sharing and premiums, supplemental health benefits) that attract more enrollees.

A plan's rebate is calculated as the difference between a plan's benchmark and bid, multiplied by a rebate percentage (which varies by the plan's quality star rating). As shown in Figure 4, lowering health care costs reduces a plan's bid, while improving quality can increase both a plan's rebate percentage and its benchmark. Either strategy results in a larger rebate and more extra benefits offered to enrollees.

Figure 4. MA coding intensity undermines the goal of plans competing on the basis of improving quality and reducing costs



Note: Rebates are only available for plans bidding below their benchmark, which is nearly all plans in 2022.

Source: MedPAC public meeting presentation, January 2022.

However, increasing coding intensity also raises a plan's benchmark, leading to a higher rebate and the ability to offer more extra benefits. Increasing coding intensity therefore may generate a competitive advantage for a plan and may crowd out efforts to improve quality or lower health care costs. At the very least, both the size of plan rebates and plan competition for more enrollees are based on a combination of all three strategies: lowering health care costs, improving quality, and increasing coding intensity.

Although the resources devoted to coding intensity offer no societal benefit, coding intensity likely increases MA enrollment as added extra benefits influence more Medicare beneficiaries to choose to enroll in an MA plan rather than FFS Medicare. For 2022, annual extra benefits in MA average nearly \$2,000 (a historic high for the sixth straight year) and account for about 15 percent of payments to MA plans. Extra benefits for reduced cost sharing may be financially beneficial to MA enrollees, but policy makers have no information about the value of the supplemental benefits that plans offer nor about beneficiaries' use of them.<sup>14</sup>

<sup>14</sup> The most commonly offered supplemental benefits in 2021 were: worldwide emergency care; routine eye exam; worldwide urgent care; fitness benefit; annual physical exam; routine hearing exam; eyewear, contacts; worldwide emergency care transportation; dental, preventive cleaning and oral exam; and eyewear, lenses and frames. See

Chapter 1 ("Rebalancing Medicare Advantage benchmark policy") of the Commission's June 2021 report for more information.

### MedPAC's approach to addressing coding intensity

The Commission strongly believes that Medicare should share in the savings associated with MA, the same standard we apply when making recommendations about payment rate updates in FFS Medicare. In March 2016, we recommended a multipronged approach that would address overpayments to MA plans, improve the equity of payments net of the coding intensity adjustment across MA contracts, and encourage competition based on improving quality and reducing costs. The recommendation, which would replace the existing mandatory minimum coding intensity adjustment, 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.

MedPAC's overall strategy is to first address the underlying causes of coding intensity. Two primary sources of coding intensity are inconsistent coding across years in FFS Medicare and MA plans' use of health risk assessments.

Most diagnoses used to set the weights in the CMS–HCC risk adjustment model come from physician or outpatient claims, which, in FFS Medicare, are paid based on procedure codes. Diagnoses that appear on these claims reflect the condition with which the patient is presenting at the time of service, and thus the provider's expectation of what they will be paid by Medicare for treating a patient with that condition. Under FFS, providers have little financial incentive to add diagnoses to the claim if they were not addressed by the provider during the encounter that is the basis for the claim. In contrast, MA plans have a significant financial incentive to code as many diagnoses during an encounter as possible and to develop the institutional infrastructure to collect and maintain these codes over time. Research has shown that many plans have responded to this incentive by pursuing multiple approaches to collect diagnoses (e.g., in-home health risk assessments, chart reviews, provider pay-for-coding incentives) and document all possible codes in each year. For MA payments to be accurate, diagnoses must be coded with the same intensity in both FFS Medicare and MA.

MA plans' use of health risk assessments also contributes to coding intensity. Beneficiaries in both MA and FFS may have health risk assessments. (In FFS, most health risk assessments are provided as part of an annual wellness visit.) However, during these assessments, MA plans are far more likely to document diagnoses that are not being actively treated. We reported in 2016 that about 30 percent of the HCCs documented through health risk assessments for MA enrollees were not treated during the year, compared with about 6 percent of diagnoses that were documented through

these assessments for FFS enrollees.<sup>15</sup> Further, our analysis and a study by the Office of Inspector General (OIG) found that use of health risk assessments varies significantly across MA contracts.<sup>16</sup>

A third source of coding intensity—which we did not address in our 2016 recommendation because the data were not available at the time—is the use of chart reviews in MA. A recent analysis from OIG indicates that use of chart reviews is a significant driver of both MA coding intensity and the variation in coding intensity across MA contracts.<sup>17</sup>

In 2017, OIG found that health risk assessments and chart reviews combined accounted for \$9.6 billion in payments to MA plans. <sup>18</sup> Based on OIG's findings, we estimate that in 2017 health risk assessments and chart reviews were responsible for about two-thirds of the greater coding intensity in MA, generating 4.6 percent of total payments to MA plans that year. <sup>19</sup>

Using two years of diagnostic data would improve the accuracy of both FFS and MA diagnostic information and would reduce year-to-year variation in documentation. We note that the 21st Century Cures Act (the Cures Act) codifies the Secretary's authority to use two years of diagnostic data in MA risk adjustment, stating that, for 2019 and subsequent years, "the Secretary may use at least two years of diagnosis data." However, CMS did not take this step in any of the rulemaking that implemented the Cures Act provisions.

At the same time, excluding diagnoses documented only through health risk assessments from risk adjustment calculations would ensure that only diagnoses that were both documented on an assessment *and* associated with medical treatment would count toward risk adjustment. We note that MA plans could still conduct health risk assessments and document diagnoses as a way to identify conditions, coordinate care, and improve outcomes. Though it was not part of our 2016 recommendations, eliminating chart reviews as a source of diagnostic data for risk adjustment would be consistent with our overall approach.

<sup>&</sup>lt;sup>15</sup> Medicare Payment Advisory Commission. 2016. *Report to the Congress: Medicare payment policy*. Washington, DC: MedPAC.

<sup>&</sup>lt;sup>16</sup> 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.

<sup>&</sup>lt;sup>17</sup> Office of Inspector General, Department of Health and Human Services. 2019. *Billions in estimated Medicare Advantage payments from chart reviews raise concerns*. OEI-03-17-00470. Washington, DC: OIG.

<sup>&</sup>lt;sup>18</sup> 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.

<sup>&</sup>lt;sup>19</sup> We estimate that MA coding intensity accounted for about 7.1 percent, or \$14.8 billion of the \$209 billion Medicare paid MA plans in 2017.

After addressing the underlying causes of coding intensity to the maximum extent possible (by using two years of FFS and MA diagnostic data in MA risk adjustment and by eliminating from risk adjustment calculations any FFS and MA diagnoses collected during health risk assessments and chart reviews that did not lead to treatment), CMS should adjust for any remaining coding intensity differences. Addressing remaining coding differences could be accomplished with an across-the-board adjustment (which is likely to be much smaller than the current adjustment), or by using a tiered approach that would group contracts into tiers of high, medium, and low coding intensity, and then apply a coding intensity adjustment based on each tier's average level of coding intensity.

#### Conclusion

Given the financial impact of coding intensity under a robust MA program on the Medicare program at large, we urge the Secretary and CMS to immediately implement a payment adjustment that fully recaptures the financial impacts of this coding and addresses the underlying causes of coding intensity. We appreciate 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 James E. Mathews, MedPAC's Executive Director, at 202-220-3700.

Sincerely,

Michael E. Chernew, Ph.D.

mucher

Chair

MC/aj