Summary of VBID Model Risk Score Additional Analytics

The enclosed memoranda document the results of the VBID model evaluation contractor's additional analyses. These analyses expand on the 2023 VBID evaluation report finding of an association in 2020 between VBID plan participation and higher beneficiary-level risk scores among enrollees eligible for VBID interventions. These additional analyses had two objectives:

1. Identify whether the overall association between the VBID model and higher enrollee risk scores was concentrated among specific VBID intervention(s) or group(s) of model participants (e.g., Dual Special Needs Plans (D-SNPs), non-SNPs); and
2. Assess which parts of the risk score calculation drove the 2023 evaluation report's observed increase in enrollee risk scores.

To the first objective, the results do not indicate that the overall risk score association with the model was attributable to specific VBID intervention(s) or group(s) of participants. Although the magnitude varied depending on the specific subset of the model that was analyzed (e.g., targeting of reduced or eliminated Part D cost-sharing), the association of the VBID model with higher risk scores among VBID-eligible enrollees persisted across the specific VBID interventions examined, and also was present for enrollees in both D-SNPs and non-SNPs that participated in the model.

To the second objective, the results of the additional analyses identified an increase in the number of Hierarchical Condition Categories (HCCs) for VBID-targeted enrollees[1], relative to comparable enrollees in plans that did not participate in the model. This included statistically significant increases in the prevalence of HCCs that were both related and unrelated to VBID interventions. An analysis of the rate at which continuously-enrolled enrollees[2] became dual eligible, which is included in the risk score calculation, did not find a substantial association between increased enrollee dual eligibility and their plan's VBID participation status.

Based on these additional findings, the association between VBID plan participation and higher enrollee risk scores was seen across subsets of the model and that the increased prevalence of HCCs drove the risk score increase associated with the VBID model.

[1] Targeted enrollee means an individual enrolled in an MA plan participating in the VBID model eligible for VBID interventions.

[2] Continuous enrollee means an individual enrolled in the same MA plan from 2019-2020.

Project Report

The Association Between the Medicare Advantage Value-Based Insurance Design Model Test and Beneficiary Risk Scores

Supplementary Analyses of 2023 Evaluation Report Findings

Christine Eibner, Michael Dworsky, Dmitry Khodyakov, Julie Lai, Erin A. Taylor, Shiyuan Zhang

RAND Health Care

PR-A1881-5
December 2024
Prepared for the Centers for Medicare & Medicaid Services

About This Report

In RAND's 2023 evaluation of the Medicare Advantage Value-Based Insurance Design (VBID) model test, we found an association between plans' participation in the VBID model and increases in beneficiary risk scores. In this report, we conduct sensitivity analyses to better understand the relationship between VBID and risk scores.

This research was funded by the Innovation Center under Research, Measurement, Assessment, Design, and Analysis Contract Number 75FCMC19D0093, Order Number 75FCMC20F0001, for which Julia Driessen is the contracting officer's representative. It was carried out within the Payment, Cost, and Coverage Program in RAND Health Care.

For more information, see www.rand.org/health-care, or contact

RAND Health Care Communications
1776 Main Street
P.O. Box 2138
Santa Monica, CA 90407-2138
(310) 393-0411, ext. 7775
RAND_Health-Care@rand.org

Acknowledgments

We thank Daniel Crespin, Jeanne Ringel, and Paul Koegel of RAND for their reviews of this document, and Leslie Canterbury for providing help with citation formatting.

Abbreviations

CHF	congestive heart failure
CI	confidence interval
CMS	Centers for Medicare & Medicaid Services
COVID-19	coronavirus disease 2019
DD	difference-in-differences
DSNP	dual eligible special needs plan
HCC	hierarchical condition category
MA	Medicare Advantage
PO	parent organization
RI	rewards and incentives
SES	socioeconomic status
VBID	Value-Based Insurance Design
WHP	wellness and health care planning

Contents

Figures and Tables

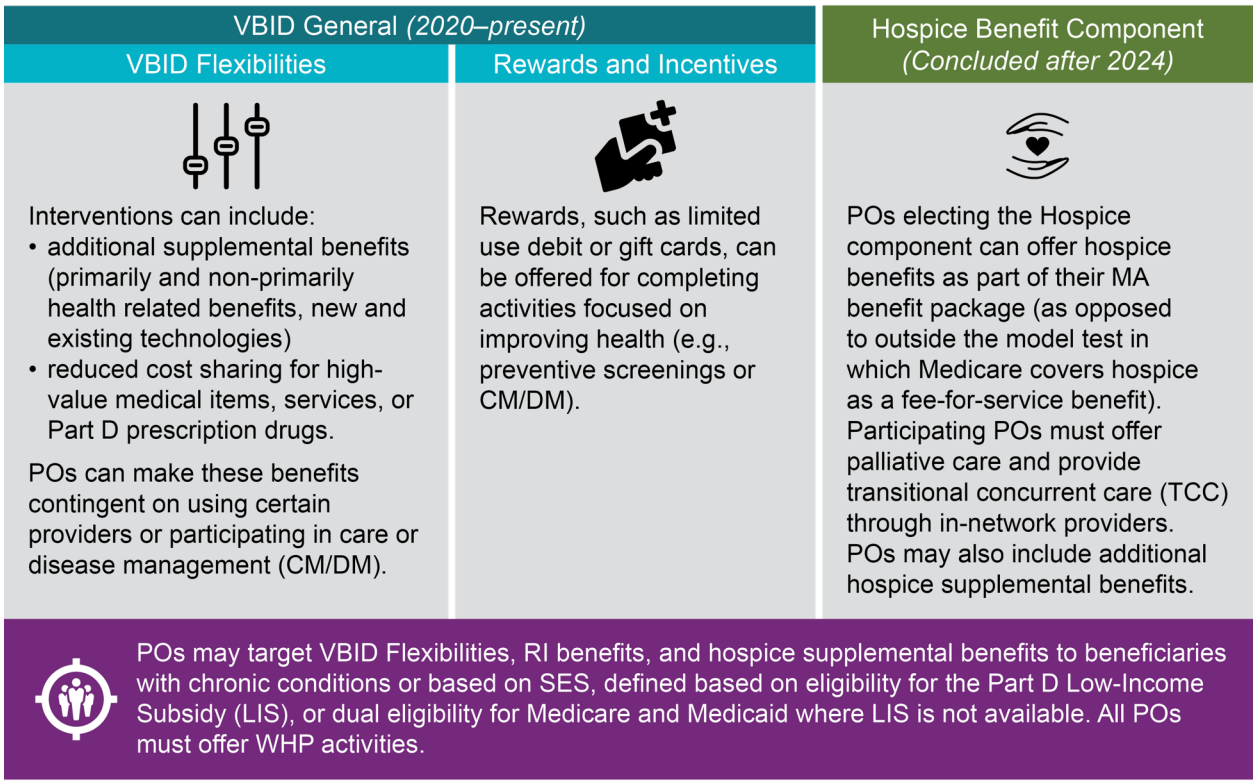
Figures

Tables

The Association Between the Medicare Advantage Value-Based Insurance Design Model Test and Beneficiary Risk Scores: Supplementary Analyses of 2023 Evaluation Report Findings

RAND researchers are conducting a multiyear evaluation of the Centers for Medicare & Medicaid Services' (CMS') Medicare Advantage (MA) Value-Based Insurance Design (VBID) model test. Initiated in 2020, the VBID General component of the VBID model test enables participating insurers to target reduced cost-sharing, additional supplemental benefits, or rewards and incentives (RI) to enrollees with one or more chronic conditions or based on socioeconomic status (SES). From 2021 to 2024, the model also allowed MA plans to offer hospice benefits directly to enrollees (outside of VBID, hospice care is "carved out" of MA and offered through Original Medicare). All model participants must offer wellness and health care planning (WHP) activities, which focus on improving awareness and availability of advance care planning, to all enrollees in their VBID-participating plans. Figure 1 shows the components of the model that existed in 2023. A Cash Rebates option that enabled VBID General plans to share MA rebates directly with beneficiaries as cash or monetary transfers was in effect in 2021 and 2022 only.

Figure 1. 2023 VBID Model Components



In our 2023 evaluation report for the VBID model (Eibner et al., 2023a), we found that plans' participation in VBID General was associated with a 0.075 point increase in average hierarchical condition category (HCC) risk scores among targeted beneficiaries in 2020 (95% confidence interval [CI]: 0.067 to 0.079), the first year of the current version of the model test. Based on diagnoses and such beneficiary characteristics as age, gender, and dual-eligibility status, *risk scores* are measures used to adjust CMS' payments to MA plans to reflect the anticipated spending of enrollees (Medicare Payment Advisory Commision, 2024). CMS pays plans by multiplying the plan bid by the risk score, so higher scores result in increased payment to plans. The risk score is normalized to 1 for a population with standard risk; the average risk score among beneficiaries targeted by 2020 VBID interventions was 1.7 (standard deviation = 1.3).

To better understand the reasons for the association between VBID General participation and beneficiary risk scores in 2020, CMS asked that we analyze whether the association between VBID and risk scores varied for dual eligible special needs plans (DSNPs) and non-SNPs. DSNPs are a specific type of MA plan that provide benefits to low-income beneficiaries who are dually eligible for Medicare and Medicaid. Differences between DSNPs and non-SNPs are of interest because, over time, an increasing share of DSNPs have joined the model test, accounting for nearly half of all VBID General plans in 2023. Additionally, CMS asked that we consider whether risk scores evolved differently depending on the type of targeting that VBID plans used (for example, targeting based on SES or chronic conditions). Risk score trends may have differed based on targeting approach because interventions varied across these approaches. For example, plans that targeted chronic conditions often implemented interventions that increased access to high-value care—and hence may have increased the likelihood that beneficiaries received diagnoses. SES-based interventions, in contrast, often focused on social needs, such as grocery cards, rather than specific health care services.

Finally, CMS asked that we analyze factors that could have contributed to the increase in risk scores. We considered whether VBID General was associated with increased diagnoses given the substantial role they play in the risk score and the possibility that they could be sensitive to insurer and provider decisionmaking (Kronick, 2017; Jung, Feldman, and Carlin, 2023; Jacobs, 2024). We also assessed whether there was a link between plans' participation in VBID General and the likelihood that a beneficiary became classified as dually eligible for Medicare and Medicaid, which factors into the risk score formula. We considered transitions to both partial dual status and full dual status, which receive different risk score weights.[1]

[1] Full dual-eligible beneficiaries qualify for Medicare and all Medicaid benefits offered by their state, while partial dual-eligible beneficiaries qualify for Medicare and a more limited set of wraparound benefits offered through Medicare Savings Programs (Centers for Medicare & Medicaid Services, 2023).

We found that

- Plans' participation in VBID General was associated with increases in targeted beneficiaries' risk scores for both non-SNPs and DSNPs, although the increase was larger in non-SNPs.
- While associations with increased risk scores were larger for beneficiaries enrolled in VBID General plans with chronic conditions targeting than they were for beneficiaries enrolled in VBID General plans with SES targeting, the associations were positive and statistically significant in both cases.
- VBID General was associated with an increase in targeted beneficiaries' total HCC-relevant diagnoses, a term we use to reflect diagnoses included in the calendar year 2020 risk score model (Centers for Medicare & Medicaid Services, 2019).
- Diagnoses with large increases in prevalence in the VBID group relative to the comparison group included conditions that were commonly targeted by VBID participants, such as congestive heart failure (CHF), diabetes with chronic complications, and rheumatoid arthritis, as well as such related conditions as vascular disease and morbid obesity. This suggests that VBID may be contributing to increases in risk scores by enabling plans to identify new or reestablish existing diagnoses.[2]
- VBID was associated with a very small increase in the probability that a beneficiary transitioned to full dual status between 2019 and 2020.

In what follows, we present the results of these analyses in more detail.

Methods Overview

Our analysis focused on beneficiaries who were eligible for their plan's VBID interventions given the targeting criteria used by the plan. We then used difference-in-differences (DD) models to compare VBID-targeted beneficiaries in participating plans with a comparison group of beneficiaries in eligible, nonparticipating plans. We weighted the comparison sample to resemble the VBID group across more than 90 characteristics using entropy-balancing weights (Hainmueller, 2012). Because our DD methodology requires a comparison of outcome trends before and after the model was implemented, we restricted our sample to beneficiaries who were enrolled in the same plan for at least one full year prior to the start of the VBID model test in 2020. The outcome measure used in our analysis of HCC risk scores was the final risk score calculated with diagnoses in the current year and used for payment in the subsequent year. We focused on year of diagnosis rather than year of payment to align the timing of the risk score diagnoses with the timing of model implementation. For example, if targeted beneficiaries in a plan that implemented VBID General in 2020 were more likely to receive diagnoses in that year because of VBID-associated utilization or coding changes, this would be reflected in the final

[2] Diagnoses must be reestablished on an annual basis to be included in risk score calculations (see Yeatts and Sangvai, 2016).

risk score used for payment in 2021. Except where noted, we derived new entropy-balancing weights for each regression and outcome presented below to ensure balance within the subgroup under consideration (for example, DSNPs, non-SNPs). For a more detailed discussion of our methodology, see Appendix C of our 2023 report (Eibner et al., 2023b).

Comparison of DSNPs and Non-SNPs

Table 1 compares the estimated association between VBID General participation and risk scores for targeted beneficiaries in DSNPs and targeted beneficiaries in non-SNPs.[3]

We found a positive association between VBID General participation and risk scores for targeted beneficiaries in both DSNPs and non-SNPs. However, the effect was smaller for DSNP enrollees than non-SNP enrollees. Specifically, VBID General was associated with a marginally significant 0.053 point (3.2%) increase in risk scores for DSNP enrollees ($p = 0.05$) and a statistically significant 0.084 point (5.0%) increase for non-SNP enrollees ($p < 0.01$). The 0.031 point difference between the coefficient for DSNPs and non-SNPs was statistically significant at the 95% confidence level (95% CI: 0.029 to 0.036).

Table 1. Association Between VBID General and Beneficiary-Level Final Risk Scores, by Plan DSNP Status, 2020

Plan Type	Coefficient (Standard Error)	95% CI	Percentage Change	Effective Sample Size
DSNPs	0.053* (0.029)	−0.0003 to 0.106	3.2%	46,041
Non-SNPs	0.084*** (0.003)	0.077 to 0.091	5.0%	345,560

NOTE: * denotes $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. Results represent DD coefficients from entropy-weighted regressions comparing VBID-targeted beneficiaries in a stable cohort with entropy-balanced comparators. The comparison group for the DSNP analysis was limited to beneficiaries in other DSNPs, and the comparison group for the non-SNP analysis excluded SNPs. Regressions were entropy-weighted across more than 90 characteristics including pre-VBID outcome trends. Weights for the DSNP analysis differed slightly from those used in our main report because we excluded characteristics that were highly related to DSNP status, such as share of plan beneficiaries that were dually eligible.

Comparison of SES and Chronic Conditions Targeting

Table 2 shows associations between VBID and targeted beneficiaries' risk scores for VBID Flexibilities plans,[4] stratified based on the approach that the plans used to target beneficiaries.

[3] The non-SNP group in this analysis excludes beneficiaries enrolled in any type of SNP, including DSNPs, institutional special needs plans and chronic condition special needs plans.

[4] VBID Flexibilities plans are a subtype of VBID General plans that may target reduced cost-sharing or additional supplemental benefits to enrollees based on chronic conditions or SES.

Although we found positive associations for both types of VBID Flexibilities plans, the association was twice as large for beneficiaries in plans that targeted their interventions based on chronic conditions as it was for beneficiaries in plans with SES targeting. This 0.06 point difference was statistically significant at the 95% confidence level (95% CI: 0.059 to 0.061).

Table 2. Association Between VBID General and Beneficiary-Level Final Risk Scores, by Targeting Approach, VBID Flexibilities Plans, 2020

Plan Type	Coefficient (Standard Error)	95% CI	Percentage Change	Effective Sample Size
Chronic conditions targeting	0.120*** (0.005)	0.110 to 0.130	6.5%	151,735
SES targeting	0.060*** (0.005)	0.049 to 0.070	4.0%	153,860

NOTE: * denotes $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. Results represent DD coefficients from entropy-weighted regressions comparing VBID-targeted beneficiaries in a stable cohort with entropy balanced comparators. Regressions were entropy-weighted across more than 90 characteristics, including pre-VBID outcome trends.

Associations with Diagnosis Counts

We considered whether VBID General was associated with an increased count of HCC diagnoses that were included in the 2020 risk score model. Each diagnosis in the risk score model is assigned a weight based on the anticipated spending impact associated with that diagnosis. Because higher-weighted diagnoses have a larger impact on beneficiaries' total risk score than lower-weighted diagnoses, we also analyzed whether VBID General was associated with changes in higher-weighted HCC diagnoses, lower-weighted HCC diagnoses, or both. We classified diagnoses that received a weight above the median value of 0.351 in the 2020 risk score model as higher-weighted and those that received a weight of 0.351 or below as lower-weighted diagnoses. Because we estimated the models using Poisson regressions, we transformed coefficients from a log scale before interpreting them. We thus focus on the implied percentage changes rather than the coefficients in our discussion. (The estimates are small, so the percentage changes are close to the coefficients.)

We found that VBID General was associated with an increase in the total number of HCC diagnoses, the total number of higher-weighted HCC diagnoses, and the total number of lower-weighted HCC diagnoses. The effect size was slightly larger for lower-weighted diagnoses relative to higher-weighted diagnoses (4.3% versus 3.9%), and the difference was statistically significant. Overall, these results imply that VBID General was associated with increased diagnoses among both lower-weighted and higher-weighted diagnoses. Although the increase was slightly larger among lower-weighted diagnoses, higher-weighted diagnoses may have a bigger impact on the overall risk score given their larger value.

Table 3. Association Between VBID General and Counts of HCC Indicators, 2020

HCC Counts	Coefficient (Standard Error)	Percentage Change	95% CI on Percentage Change	Effective Sample Size
Total number of HCCs	0.042*** (0.002)	4.3%	3.9% to 4.6%	807,6401
Total number of higher-weighted HCCs (weight > 0.351 in the 2020 risk score model)	0.038*** (0.003)	3.9%	3.3% to 4.6%	807,539
Total number of lower-weighted HCCs (weight ≤ 0.351 in the 2020 risk score model)	0.043*** (0.002)	4.4%	4.1% to 4.8%	807,693

NOTE: * denotes $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. Results represent DD coefficients from entropy-weighted regressions comparing VBID-targeted beneficiaries in a stable cohort to entropy-balanced comparators. Regressions were entropy-weighted across more than 90 characteristics, including pre-VBID outcome trends. Because the dependent variables were counts, we estimated the models using Poisson regression, which are interpretable on a long scale.

We further explored whether the change in HCC counts differed for DSNPs and non-SNPs (Table 4). These analyses apply the same restrictions for the comparison group as described above; namely, the DSNP comparison is limited to other DSNPs, and the non-SNP comparison group is limited to non-SNPs.

For DSNPs, the association between VBID and diagnosis counts was present only for diagnoses with a below-median weight in the 2020 risk score model (4.8%, 95% CI: 1.7% to 8.1%). The estimated association for higher-weighted HCCs among DSNPs, in contrast, was statistically insignificant and close to zero (−0.2%, 95% CI: −7.3% to 7.4%). For non-SNPs, we estimated a positive association between VBID and diagnosis counts for all HCCs, higher-weighted HCCs, and lower-weighted HCCs. As with the main model shown in Table 3, we found that the relationship between VBID and diagnosis counts for non-SNPs was slightly larger for counts of lower-weighted HCCs compared with counts of higher-weighted HCCs. While the difference between higher-weighted and lower-weighted HCCs was statistically significant for both DSNPs and non-SNPs ($p < 0.01$ in both cases), lower-weighted diagnoses appear to drive the increase in risk scores among DSNPs, whereas both lower and higher-weighted diagnoses contribute to risk score changes for non-SNPs.

Table 4. Association Between VBID General and Counts of HCC Indicators, 2020, DSNPs Compared with Non-SNPs

	DSNPS				Non-SNPs			
HCC Counts	Coef. (SE)	Percentage Change	95% CI on Percentage Change	ESS	Coef. (SE)	Percentage Change	95% CI on Percentage Change	ESS
Total HCCs	0.033 (0.021)	3.4%	−0.6% to 7.5%	46,066	0.042*** (0.002)	4.3%	3.9% to 4.8%	345,376
Higher-weighted HCCs	−0.002 (0.039)	−0.2%	−7.3% to 7.4%	46,065	0.041*** (0.004)	4.2%	3.4% to 5.0%	345,380
Lower-weighted HCCs	0.047*** (0.017)	4.8%	1.7% to 8.1%	46,065	0.044*** (0.002)	4.5%	4.1% to 4.9%	345,361

NOTE: * denotes $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. Results represent DD coefficients from entropy-weighted regressions comparing VBID-targeted beneficiaries in a stable cohort with entropy-balanced comparators. Regressions were entropy-weighted across more than 90 characteristics, including pre-VBID outcome trends. Because the dependent variables were counts, we estimated the models using Poisson regression. Coef. = coefficient; ESS = effective sample size; SE = standard error.

Analysis of Specific Diagnoses

To better understand which specific diagnoses contributed to the increase in HCCs, we conducted a descriptive analysis in which we estimated the 2019 to 2020 change in the share of our sample that had each HCC diagnosis, for both the VBID General and the comparison groups. We then tested whether the difference between the two changes was statistically significant.

This methodology is a form of DD analysis, similar to the approach used in our regressions. However, for this descriptive approach, we adjusted the comparison group using a single set of weights that included a comprehensive set of plan and beneficiary characteristics rather than estimating unique weights for each outcome. We implemented this descriptive approach and used a single set of weights because it greatly reduced the computational resources required to conduct the analysis, compared with calculating new weights and running a separate regression for each HCC diagnosis. The primary drawback of the descriptive approach is that, although we controlled for pre-period trends in average risk scores in the VBID and comparison groups, we cannot rule out the possibility that pre-period trends for specific diagnoses differed in the VBID and comparison groups. An additional drawback of the descriptive approach is that we cannot adjust for regression controls used in the main models, including COVID-19 case rates and plans' participation in other CMS initiatives, such as Uniformity Flexibility. Despite these limitations, our approach enables us to assess how HCC diagnoses among targeted beneficiaries in VBID plans changed relative to similar beneficiaries in comparison plans, holding constant a large range of beneficiary, plan, and community-level characteristics.

Table 6 at the end of this report shows the results of this analysis. For the vast majority of the HCCs, we found an increase in the share of diagnosed beneficiaries among the VBID General

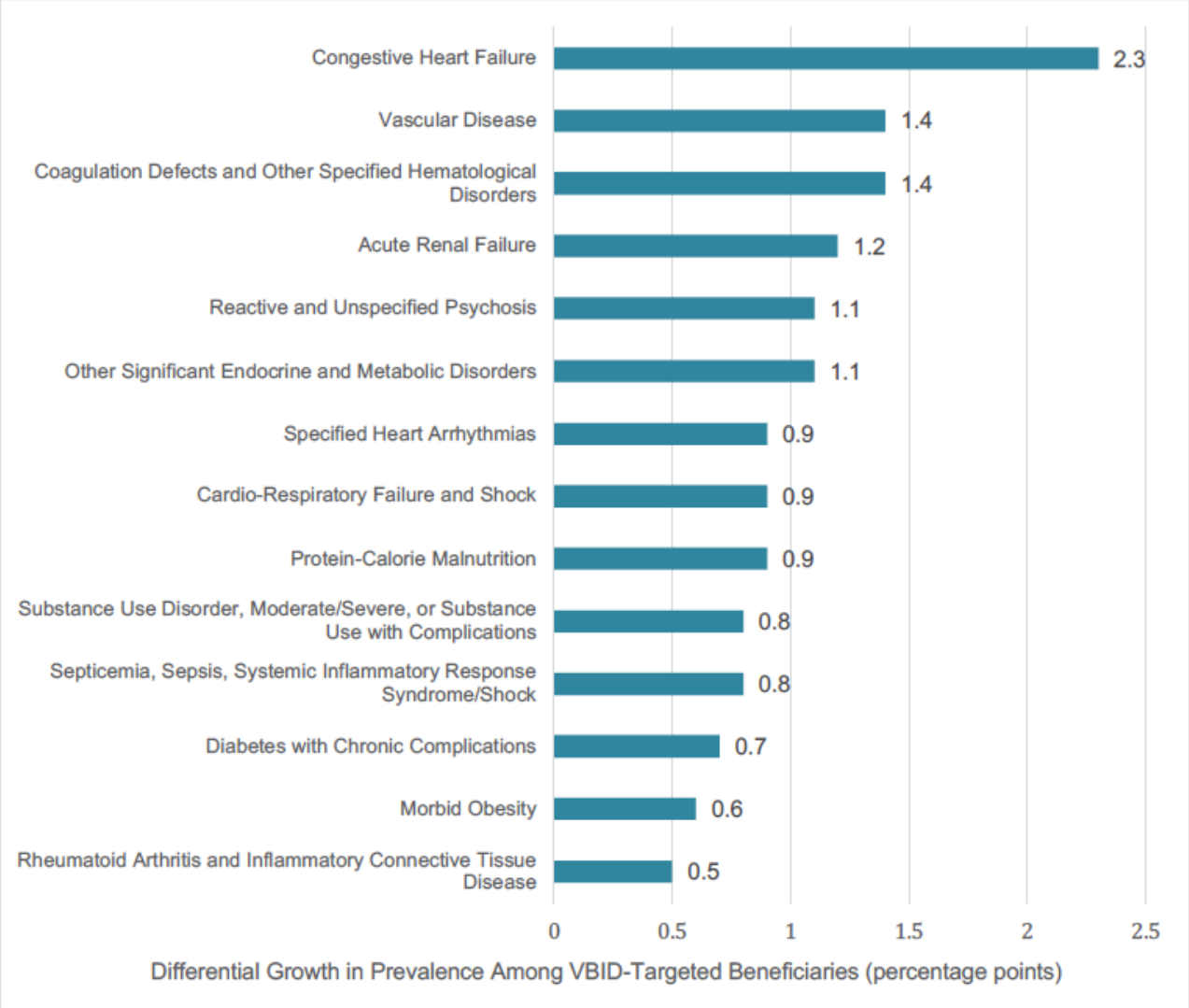
group relative to the comparison group. Because of the large number of beneficiaries in our analysis, these descriptive DD estimates were nearly all statistically significant, even when the magnitude of the estimate was small.[5] Because most of these changes were small in absolute value, we focused on estimates that are at least 0.5 percentage points.

Figure 2 shows the estimated increase in the prevalence of the diagnoses in the VBID General group relative to the comparison group for these conditions. For example, the top line of the figure shows that CHF prevalence in the VBID group increased by 2.3 more percentage points between 2019 and 2020 among VBID-targeted beneficiaries than among entropy-balanced comparison beneficiaries. (The detailed results in Table 6 clarify that CHF prevalence increased in both groups, by 2.7 percentage points in the VBID group and by 0.4 percentage points in the comparison group, for a DD of 2.3 percentage points.)

Several diagnoses that had meaningful increases in the VBID group relative to the comparison group—including CHF, diabetes with chronic complications, and rheumatoid arthritis—were conditions that plans directly targeted with their VBID interventions. To target beneficiaries based on chronic conditions, insurers had to design approaches to identify beneficiaries with these conditions, which could have helped them code diagnoses for these conditions more comprehensively and reestablish diagnoses on an annual basis. Other conditions shown in the figure, such as vascular disease and morbid obesity, are related to VBID-targeted conditions and may have been identified or coded during such VBID activities as care management. Still other diagnoses in the figure may be related to the increase in hospitalizations among VBID-targeted beneficiaries that we estimated in the 2023 evaluation report (Eibner et al., 2023a) or with VBID's increased focus on addressing social determinants of health.

[5] All but two of the differences were statistically significant, including differences that were extremely small and not meaningfully different from zero. The two differences that were not statistically significant were for HCC136 (chronic kidney disease, stage 5) and HCC79 (Parkinson's disease). Only two results were negative; for pneumococcal pneumonia (HCC115) and diabetes without complications (HCC19).

Figure 2. DD Estimates (Descriptive Approach) for Growth in Prevalence of Specific Diagnoses Among VBID-Targeted Relative to Comparison Beneficiaries, 2019–2020



NOTE: The DD estimates reflect the 2019 to 2020 change in prevalence in the treatment group minus the 2019 to 2020 change in prevalence in the entropy-weighted comparison groups, measured in percentage points. The entropy weights adjust for over 90 plan and beneficiary-level characteristics, including trends in risk scores. This figure shows all DD estimates that were at least 0.5 percentage points. Table 6 at the end of this report shows results for all conditions analyzed. All results shown in the figure were statistically significant ($p \leq 0.01$).

Analysis of Dual Status

Our DD regression analyses controlled for fixed characteristics, such as race and sex, that do not change over time, and age, which changes on a predictable basis. However, dual status also contributes to the risk score. Although our regression analyses also controlled for dual status at baseline (that is, whether a beneficiary was dually eligible for Medicaid and Medicare in 2019), it is possible that beneficiaries in VBID plans were more likely to transition to dual status between 2019 and 2020. For example, if insurers offered VBID in plans in which beneficiaries

had significant health-related social needs, non-dually eligible beneficiaries in VBID plans might have had a higher probability of becoming eligible for Medicaid between 2019 and 2020 (for example, because of income instability) than non-dually eligible beneficiaries in comparison plans.

In Table 5, we show a descriptive DD analysis that assessed change in dual status among VBID-targeted beneficiaries relative to comparators using the same approach as described above for the analyses of specific diagnoses.

In both 2019 and 2020, VBID-targeted beneficiaries were more likely to have full or partial dual status than comparison beneficiaries. However, very few beneficiaries transitioned into or out of dual status over this period. Beneficiaries in both the VBID and the comparison groups experienced a net decline in the probability of being categorized as *non-dual* by less than 1 percentage point. Similarly, for both VBID and comparison beneficiaries, there was a slight decline in the probability of being categorized as *partial dual* and a slight increase in the probability of being categorized as *full dual*.

Although the percentage point changes in all cases were larger in absolute value for the VBID group than for the weighted comparison group, these differences were small in magnitude. For example, the results indicate that between 2019 and 2020, VBID beneficiaries were 0.3 percentage points more likely to transition to full dual status than comparison beneficiaries and 0.2 percentage points less likely to retain non-dual status. While these results are statistically significant, we estimate that the 0.3 percentage point increase in the probability of having full dual status would increase the average risk score by 0.001 points,[6] less than 2% of the 0.075 point increase in risk scores associated with VBID.

[6] To calculate this impact, we computed the average risk score among beneficiaries in our sample and then estimated how much this would change if we changed the proportion of beneficiaries with full, partial, and non-dual status based on the proportions shown in Table 5.

Table 5. DD Estimates of the Impact of VBID on Targeted Beneficiaries' Dual Status on Risk Scores, Measure Year 2020

Beneficiary Group	2019	2020	Difference
VBID, full dual	22.2%	22.9%	0.7%
Weighted comparison, full dual	19.0%	19.3%	0.4%
			0.3%***[a]
VBID, partial dual	17.3%	17.1%	−0.2%
Weighted comparison, partial dual	15.2%	15.1%	−0.1%
			−0.1%**[a]
VBID, non-dual	60.4%	60.0%	−0.4%
Weighted comparison, non-dual	65.8%	65.5%	−0.3%
			−0.2%**[a]

NOTE: The analysis includes 165,651 VBID-targeted beneficiaries and 8,045,903 comparison beneficiaries (unweighted) who were stably enrolled in their plans for at least one year prior to the start of the model test. Asterisks imply that the change for the VBID group was statistically significantly different from change for the comparison group, with $p < 0.01$ (**) or $p < 0.001$ (***).
[a] DD estimates are shown in bold.

Limitations

In interpreting the results of this report, it is important to keep in mind that our intervention group is limited to VBID-targeted beneficiaries who were in the same plan for at least one year prior to the start of the VBID model test. This group represents only a subset of beneficiaries who are enrolled in each VBID-participating plan. It is possible that the trend that we observe for our intervention sample differed from the overall trend in risk scores among all beneficiaries enrolled in VBID plans. Furthermore, our convention in this analysis has been to link beneficiaries to risk scores based on the year of diagnosis rather than the year used for payment. That is, the *2020 risk score* in this analysis reflects the risk scores that were calculated based on 2020 diagnoses and used for payment in 2021. Focusing on the year of diagnosis makes sense given the possibility that VBID could affect diagnoses through increased interactions with the health system, changes in coding practices, or changes in beneficiaries' health status. However, these differences in timing should be considered when comparing results presented in this report to plan-level risk scores that are tracked based on the year used for payment.

In addition, we only analyzed a single year of data after VBID implementation; a longer time series may be required to understand whether these results persist. Such results may be presented in future evaluation reports. Furthermore, health care utilization during the VBID implementation year that we analyzed (2020) was affected by the COVID-19 pandemic. Although the pandemic affected both the treatment and comparison groups and we controlled for COVID-19 case rates in our regressions, it is possible that plan or community-level differences in response to the pandemic affected our results. Finally, because the VBID model test was

voluntary, plans that entered the model were selectively different from plans that did not. To address this concern, our entropy-balancing algorithm adjusted for a variety of plan and beneficiary characteristics that may differ between the treatment and comparison group, including pre-VBID trends in risk scores. Controlling for pre-VBID risk score trends reduces the possibility that our results are driven by preexisting differences in coding strategies in treatment plans relative to comparison plans. However, we cannot rule out the possibility that unmodeled differences between the VBID and comparison groups may be partly responsible for our results.

Discussion

In this report, we presented stratified analyses to better contextualize the positive association between VBID General implementation and targeted beneficiaries' risk scores that we found in our 2023 evaluation report (Eibner et al., 2023a). Our sensitivity analyses indicate that the association between VBID General and risk scores was present among targeted beneficiaries in both DSNPs and non-SNPs, although the effects were larger for beneficiaries in non-SNPs. Similarly, our analyses found that VBID was associated with an increase in risk scores in VBID Flexibilities plans with both SES and chronic conditions targeting; however, the increase was larger in plans with chronic conditions targeting. Non-SNPs were more likely than DSNPs to target beneficiaries based on chronic conditions, so the findings are likely related.

Our results also suggest that the risk score increases were driven at least in part by increases in diagnoses, especially diagnoses that get lower weights in the risk score calculation. The increase in diagnoses is consistent with findings from our interviews with insurers, which indicated that some of them used certain elements of the VBID model test, such as the required WHP activities, as an opportunity to ensure that diagnoses were comprehensively coded. For example, one insurer representative stated that the "annual wellness visit is when we collect all of conditions, so that improves risk scores. It makes the risk scores accurate. . . . We also offer in-home visits for risk adjustment and some quality-gap closing." Although insurers generally told us that they used existing tools to implement VBID's WHP requirements because health care planning and annual wellness visits are not new and all plans must offer them, some VBID participants placed renewed attention on encouraging beneficiaries to complete annual wellness visits and advance directives by providing them or their primary health care physicians with financial incentives.

Some insurers also emphasized the role of in-home visits to make it easier for beneficiaries to complete health risk assessments and encouraged their clinical teams to place additional focus on beneficiaries from VBID-participants plans. A representative of one insurer that offered VBID in a DSNP said that

> the wellness and healthcare planning allows us an uptick [in risk scores], certainly similar to the programs that we have, but with the clinical team having a little bit stronger focus because of doing things slightly differently [for VBID-

> participating plans]. We think that that might have helped to contribute to better coding for these members.

Our descriptive analyses found that diagnoses with large increases in the VBID General group relative to the comparison group included such VBID-targeted conditions as CHF and diabetes with chronic complications. To offer benefits to people with these conditions, insurers participating in VBID needed to develop approaches to identify and track eligible beneficiaries. It is possible that these beneficiary identification approaches helped insurers uncover or reestablish diagnoses that previously might not have been properly documented. Several insurers focused their RI interventions on rewarding beneficiaries who completed screenings or participated in care management, which could have also helped them identify new diagnoses.

Although we found a slight increase in the chance that VBID beneficiaries were classified as fully dual-eligible for Medicaid relative to comparators, this difference was small. We estimate that this shift could explain less than 2% of the 0.075 point risk score increase (95% CI: 0.067 to 0.079) identified in the main analysis.

Higher risk scores may more accurately reflect beneficiaries' diagnoses and health care needs, but higher risk scores also increase payment to MA plans, potentially increasing costs to CMS. Future work is needed to understand whether these risk score changes were necessary to improve or maintain beneficiaries' health.

Table 6. DD Estimates (Descriptive Approach) for Growth in Prevalence of Specific HCC Diagnoses Between 2019 and 2020

HCC	Description Label	Share of Tx Beneficiaries with Diagnosis, 2019	Share of Tx Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, Treatment	Weighted Share of EBC Beneficiaries with Diagnosis, 2019	Weighted Share of EBC Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, EBC	Difference in Difference
HCC85	Congestive Heart Failure	0.285	0.312	0.027	0.214	0.218	0.004	0.023
HCC48	Coagulation Defects and Other Specified Hematological Disorders	0.116	0.141	0.025	0.119	0.129	0.011	0.014
HCC108	Vascular Disease	0.374	0.393	0.019	0.332	0.337	0.005	0.014
HCC135	Acute Renal Failure	0.075	0.081	0.006	0.074	0.068	−0.006	0.012
HCC23	Other Significant Endocrine and Metabolic Disorders	0.07	0.086	0.016	0.082	0.087	0.005	0.011
HCC58	Reactive and Unspecified Psychosis	0.237	0.252	0.015	0.222	0.226	0.004	0.011
HCC21	Protein-Calorie Malnutrition	0.035	0.044	0.009	0.035	0.035	0	0.009
HCC84	Cardio-Respiratory Failure and Shock	0.071	0.085	0.013	0.057	0.061	0.004	0.009
HCC96	Specified Heart Arrhythmias	0.193	0.208	0.016	0.164	0.171	0.007	0.009
HCC2	Septicemia, Sepsis, Systemic Inflammatory Response Syndrome/Shock	0.036	0.039	0.003	0.041	0.036	−0.005	0.008
HCC55	Substance Use Disorder, Moderate/Severe, or Substance Use with Complications	0.116	0.121	0.005	0.097	0.094	−0.003	0.008
HCC18	Diabetes with Chronic Complications	0.353	0.359	0.006	0.321	0.32	−0.001	0.007
HCC22	Morbid Obesity	0.183	0.183	−0.001	0.167	0.16	−0.007	0.006

HCC	Description Label	Share of Tx Beneficiaries with Diagnosis, 2019	Share of Tx Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, Treatment	Weighted Share of EBC Beneficiaries with Diagnosis, 2019	Weighted Share of EBC Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, EBC	Difference in Difference
HCC40	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	0.116	0.115	−0.001	0.113	0.107	−0.006	0.005
HCC111	Chronic Obstructive Pulmonary Disease	0.491	0.491	0	0.247	0.243	−0.004	0.004
HCC114	Aspiration and Specified Bacterial Pneumonias	0.011	0.013	0.002	0.012	0.01	−0.002	0.004
HCC8	Metastatic Cancer and Acute Leukemia	0.011	0.015	0.004	0.017	0.018	0.001	0.003
HCC33	Intestinal Obstruction/Perforation	0.017	0.017	0	0.019	0.016	−0.003	0.003
HCC39	Bone/Joint/Muscle Infections/Necrosis	0.015	0.016	0.001	0.016	0.014	−0.002	0.003
HCC88	Angina Pectoris	0.094	0.097	0.003	0.069	0.069	−0.001	0.003
HCC103	Hemiplegia/Hemiparesis	0.029	0.031	0.002	0.032	0.031	−0.002	0.003
HCC161	Chronic Ulcer of Skin, Except Pressure	0.028	0.03	0.002	0.027	0.025	−0.001	0.003
HCC176	Complications of Specified Implanted Device or Graft	0.018	0.02	0.001	0.021	0.019	−0.002	0.003
HCC72	Spinal Cord Disorders/Injuries	0.013	0.014	0.001	0.011	0.011	0	0.002
HCC75	Myasthenia Gravis/Myoneural Disorders and Guillain-Barre Syndrome/Inflammatory and Toxic Neuropathy	0.028	0.033	0.005	0.038	0.041	0.003	0.002
HCC82	Respirator Dependence/Tracheostomy Status	0.005	0.005	0	0.008	0.006	−0.002	0.002

HCC	Description Label	Share of Tx Beneficiaries with Diagnosis, 2019	Share of Tx Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, Treatment	Weighted Share of EBC Beneficiaries with Diagnosis, 2019	Weighted Share of EBC Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, EBC	Difference in Difference
HCC86	Acute Myocardial Infarction	0.03	0.029	0	0.024	0.022	−0.002	0.002
HCC100	Ischemic or Unspecified Stroke	0.04	0.04	0.001	0.035	0.033	−0.002	0.002
HCC106	Atherosclerosis of the Extremities with Ulceration or Gangrene	0.009	0.01	0.001	0.011	0.01	−0.001	0.002
HCC107	Vascular Disease with Complications	0.031	0.032	0.001	0.026	0.026	−0.001	0.002
HCC158	Pressure Ulcer of Skin with Full Thickness Skin Loss	0.005	0.007	0.002	0.006	0.006	0	0.002
HCC169	Vertebral Fractures without Spinal Cord Injury	0.013	0.014	0.001	0.012	0.011	−0.001	0.002
HCC170	Hip Fracture/Dislocation	0.01	0.012	0.002	0.01	0.01	0	0.002
HCC188	Artificial Openings for Feeding or Elimination	0.012	0.014	0.002	0.015	0.015	0	0.002
HCC6	Opportunistic Infections	0.005	0.005	0	0.004	0.004	−0.001	0.001
HCC9	Lung and Other Severe Cancers	0.018	0.02	0.002	0.014	0.015	0.001	0.001
HCC10	Lymphoma and Other Cancers	0.014	0.014	0.001	0.015	0.015	0	0.001
HCC11	Colorectal, Bladder, and Other Cancers	0.022	0.023	0	0.019	0.019	0	0.001
HCC12	Breast, Prostate, and Other Cancers and Tumors	0.056	0.055	−0.001	0.056	0.054	−0.002	0.001
HCC27	End-Stage Liver Disease	0.007	0.008	0.001	0.009	0.008	0	0.001
HCC28	Cirrhosis of Liver	0.012	0.013	0.001	0.011	0.011	0	0.001
HCC46	Severe Hematological Disorders	0.006	0.006	0	0.007	0.007	−0.001	0.001

HCC	Description Label	Share of Tx Beneficiaries with Diagnosis, 2019	Share of Tx Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, Treatment	Weighted Share of EBC Beneficiaries with Diagnosis, 2019	Weighted Share of EBC Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, EBC	Difference in Difference
HCC47	Disorders of Immunity	0.02	0.026	0.006	0.026	0.031	0.005	0.001
HCC54	Substance Use with Psychotic Complications	0.003	0.003	0	0.003	0.002	0	0.001
HCC57	Schizophrenia	0.024	0.024	0	0.025	0.024	−0.001	0.001
HCC79	Seizure Disorders and Convulsions	0.042	0.042	0	0.045	0.043	−0.002	0.001
HCC80	Coma, Brain Compression/Anoxic Damage	0.004	0.004	0.001	0.005	0.004	−0.001	0.001
HCC87	Unstable Angina and Other Acute Ischemic Heart Disease	0.022	0.02	−0.001	0.019	0.016	−0.003	0.001
HCC99	Intracranial Hemorrhage	0.005	0.006	0.001	0.006	0.005	0	0.001
HCC124	Exudative Macular Degeneration	0.017	0.018	0.001	0.016	0.016	0.001	0.001
HCC134	Dialysis Status	0.003	0.006	0.003	0.003	0.005	0.002	0.001
HCC167	Major Head Injury	0.007	0.008	0	0.008	0.007	0	0.001
HCC173	Traumatic Amputations and Complications	0.003	0.003	0.001	0.003	0.003	0	0.001
HCC1	HIV/AIDS	0.006	0.006	0	0.007	0.007	0	0
HCC17	Diabetes with Acute Complications	0.008	0.008	0	0.007	0.007	−0.001	0
HCC29	Chronic Hepatitis	0.015	0.014	−0.001	0.015	0.013	−0.002	0
HCC34	Chronic Pancreatitis	0.006	0.006	0	0.006	0.005	0	0
HCC35	Inflammatory Bowel Disease	0.015	0.014	−0.001	0.014	0.013	−0.001	0
HCC70	Quadriplegia	0.001	0.002	0.001	0.003	0.003	0	0

HCC	Description Label	Share of Tx Beneficiaries with Diagnosis, 2019	Share of Tx Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, Treatment	Weighted Share of EBC Beneficiaries with Diagnosis, 2019	Weighted Share of EBC Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, EBC	Difference in Difference
HCC71	Paraplegia	0.002	0.003	0	0.004	0.004	0	0
HCC73	Amyotrophic Lateral Sclerosis and Other Motor Neuron Disease	0	0	0	0.001	0	0	0
HCC74	Cerebral Palsy	0.003	0.003	0	0.003	0.003	0	0
HCC76	Muscular Dystrophy	0.001	0.001	0	0.001	0.001	0	0
HCC77	Multiple Sclerosis	0.007	0.007	0	0.007	0.007	0	0
HCC78	Parkinson's and Huntington's Diseases	0.014	0.015	0.001	0.016	0.017	0.001	0[ns]
HCC83	Respiratory Arrest	0	0	0	0	0	0	0
HCC104	Monoplegia, Other Paralytic Syndromes	0.004	0.004	0	0.003	0.003	0	0
HCC110	Cystic Fibrosis	0	0	0	0	0	0	0
HCC112	Fibrosis of Lung and Other Chronic Lung Disorders	0.01	0.012	0.001	0.016	0.017	0.001	0
HCC122	Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	0.016	0.015	−0.001	0.014	0.013	−0.001	0
HCC136	Chronic Kidney Disease, Stage 5	0.003	0.003	0	0.004	0.004	0	0[ns]
HCC137	Chronic Kidney Disease, Severe (Stage 4)	0.016	0.018	0.003	0.014	0.016	0.002	0
HCC157	Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	0.002	0.002	0.001	0.002	0.003	0	0
HCC162	Severe Skin Burn or Condition	0	0	0	0	0	0	0
HCC166	Severe Head Injury	0	0	0	0	0	0	0

HCC	Description Label	Share of Tx Beneficiaries with Diagnosis, 2019	Share of Tx Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, Treatment	Weighted Share of EBC Beneficiaries with Diagnosis, 2019	Weighted Share of EBC Beneficiaries with Diagnosis, 2020	2019 to 2020 Difference, EBC	Difference in Difference
HCC186	Major Organ Transplant or Replacement Status	0.004	0.004	0	0.005	0.004	0	0
HCC189	Amputation Status, Lower Limb/Amputation Complications	0.012	0.013	0.001	0.012	0.012	0.001	0
HCC19	Diabetes without Complication	0.082	0.079	−0.002	0.068	0.067	−0.001	−0.001
HCC115	Pneumococcal Pneumonia, Empyema, Lung Abscess	0.02	0.011	−0.008	0.015	0.008	−0.007	−0.001

NOTE: Tx refers to the treatment group (the stable cohort of VBID-targeted beneficiaries), and EBC refers to the entropy-balanced comparison group. HCCs are ranked in descending order by the size of the DD estimate in the last column. Beneficiaries in the EBC group are weighted using entropy-balancing weights that adjusted for over 90 characteristics, including trends over time in the total number of HCCs recorded. All results were statistically significant at the 1% level ($p < 0.01$), including differences that could not be meaningfully differentiated from zero, except HCC78 and HCC136, which are labeled with the subscript “ns” in the final column.

References

Centers for Medicare & Medicaid Services, "Announcement of Calendar Year (CY) 2020 Medicare Advantage Capitation Rates and Medicare Advantage and Part D Payment Policies and Final Call Letter," April 1, 2019. As of November 21, 2024: https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/Announcement2020.pdf

Centers for Medicare & Medicaid Services, "People Dually Eligible for Medicare and Medicad," fact sheet, March 2023. As of November 21, 2024: https://www.cms.gov/Medicare-Medicaid-Coordination/Medicare-and-Medicaid-Coordination/Medicare-Medicaid-Coordination-Office/Downloads/MMCO_Factsheet.pdf

Eibner, Christine, Dmitry Khodyakov, Erin Audrey Taylor, Denis Agniel, Rebecca Anhang Price, Julia Bandini, Marika Booth, Lane F. Burgette, Christine Buttorff, Catherine C. Cohen, Stephanie Dellva, Michael Dworsky, Natalie Ernecoff, Priya Gandhi, Alice Y. Kim, Julie Lai, Monique Martineau, Nabeel Qureshi, Jessica Randazzo, Afshin Rastegar, Lucy B. Schulson, Daniel Schwam, Joan M. Teno, Anagha Alka Tolpadi, Asa Wilks, and Shiyuan Zhang, *Evaluation of Phase II of the Medicare Advantage Value-Based Insurance Design Model Test: First Three Years of Implementation (2020–2022)*, RAND Corporation, PR-A1881-2, 2023a. As of November 21, 2024: https://www.cms.gov/priorities/innovation/data-and-reports/2023/vbid-2nd-eval-report

Eibner, Christine, Dmitry Khodyakov, Erin Audrey Taylor, Denis Agniel, Rebecca Anhang Price, Julia Bandini, Marika Booth, Lane F. Burgette, Christine Buttorff, Catherine C. Cohen, Stephanie Dellva, Michael Dworsky, Natalie C. Ernecoff, Priya Gandhi, Alice Y. Kim, Julie Lai, Monique Martineau, Nabeel Qureshi, Jessica Randazzo, Afshin Rastegar, Lucy Schulson, Daniel Schwam, Joan M. Teno, Anagha Alka Tolpadi, Asa Wilks, and Shiyuan Zhang, *Appendices: Evaluation of Phase II of the Medicare Advantage Value-Based Insurance Design Model Test: First Three Years of Implementation (2020–2022)*, 2023b. As of November 21, 2024: https://www.cms.gov/priorities/innovation/data-and-reports/2023/vbid-2nd-eval-report-app

Hainmueller, Jens, "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis*, Vol. 20, No. 1, Winter 2012.

Jacobs, Paul D., "In-Home Health Risk Assessments and Chart Reviews Contribute to Coding Intensity in Medicare Advantage," *Health Affairs*, Vol. 43, No. 7, July 2024.

Jung, Jeah, Roger Feldman, and Caroline Carlin, "Coding Intensity Through Health Risk Assessments and Chart Reviews in Medicare Advantage: Does It Explain Resource Use?" *Medical Care Research and Review*, Vol. 80, No. 6, December 2023.

Kronick, Richard, "Projected Coding Intensity in Medicare Advantage Could Increase Medicare Spending by $200 Billion over Ten Years," *Health Affairs*, Vol. 36, No. 2, February 2017.

Medicare Payment Advisory Commision, "Medicare Advantage Program Payment System," revised October 2024. As of November 21, 2024: https://www.medpac.gov/wp-content/uploads/2024/10/MedPAC_Payment_Basics_24_MA_FINAL_SEC.pdf

Yeatts, John P. and Devdutta G. Sangvai, "HCC Coding, Risk Adjustment, and Physician Income: What You Need to Know," *Family Practice Management*, Vol. 23, No. 5, 2016.