Evaluation of the Million Hearts CVD Risk Reduction Model

Study Protocol and Statistical Analysis Plan

DRAFT (FEBRUARY 10, 2019)

CONTENTS

CONTENTS 2

I. Background and Objectives of the Million Hearts® Cardiovascular Disease Risk Reduction Model Evaluation 3

A. Background 3

B. Objectives of the Million Hearts® CVD model impact evaluation 4

II. Data Sources 5

A. Primary data collection 5

1. Provider survey 5

B. Secondary data 6

1. Million Hearts Data Registry 6

2. Medicare FFS and Part D claims and enrollment data 6

3. Other CMS data 7

III. Study design 8

IV. Outcomes 8

V. Regression modeling 11

A. Frequentist and Bayesian repeated outcome models 11

B. Time-to-event analysis 13

C. Subgroup analyses 14

ADDENDUM, MAY 2023 15

REFERENCES 16

I. BACKGROUND AND OBJECTIVES OF THE MILLION HEARTS® CARDIOVASCULAR DISEASE RISK REDUCTION MODEL EVALUATION

A. Background

Despite significant progress over the past 20 years, cardiovascular disease (CVD) remains the leading cause of death and disability in the United States, costing an estimated $450 billion in health care spending and lost productivity each year (Centers for Disease Control and Prevention [CDC] 2012). The primary risk factors for CVD (high blood pressure, high cholesterol, smoking, type 2 diabetes, and obesity) can be treated effectively and inexpensively. If these risk factors were well controlled through behavioral modification or treatment, CDC estimates that the risk for death from heart attacks and strokes in the United States would fall by more than half (CDC 2012).

In January 2017, the Centers for Medicare & Medicaid Services (CMS) launched the Million Hearts Cardiovascular Disease (CVD) Risk Reduction model, designed to reduce heart attacks and strokes among Medicare fee-for-service (FFS) beneficiaries. Medicare's current FFS payment system does not reward providers for developing and implementing innovative approaches for preventing chronic illnesses such as CVD. The Million Hearts CVD model encourages innovation by offering providers supports and financial incentives to assess and reduce the 10-year predicted risk of heart attack and stroke among their Medicare FFS beneficiaries.

Through this model, CMS is testing this core question: Do the supports and financial incentives offered to organizations under the Million Hearts CVD model reduce the number of CVD events (heart attacks and strokes) and/or total cost of care for their Medicare FFS beneficiaries? If the Million Hearts CVD model improves care quality while reducing Medicare spending at least enough to offset model payments, CMS could expand the model to Medicare FFS beneficiaries more broadly. The model could also pave the way for other value-based payment approaches to preventing chronic illnesses (Sanghavi and Conway 2015).

CMS is testing the Million Hearts CVD model in a rigorous five-year randomized trial and has randomized over 400 organizations[1] throughout the country, assigning half to the intervention and half to a control group. The participating organizations reflect a range of specialty types (for example, primary care and cardiology), ownership structures (independent and health system-owned), and locations (urban and rural). The intervention organizations are expected to do the following:

- Risk stratify all of their eligible Medicare FFS beneficiaries, using the American College of Cardiology/American Heart Association (ACC/AHA) calculator to estimate each eligible beneficiary's risk of having a heart attack or stroke over the next 10 years. Beneficiaries are eligible if they are ages 40–79 as of enrollment in the program, have not had a heart attack

[1] Participating organizations are typically primary or specialty practices (or groups of practices), although some participating organizations are hospitals or health systems.

or stroke, and meet other inclusion criteria.[2] Beneficiaries with a CVD risk exceeding 30 percent are considered high risk, whereas those with a risk from 15–30 percent are medium risk. All others are low risk.

- Provide cardiovascular care management to high-risk beneficiaries. This includes (1) shared decision making and individual risk modification planning—that is, helping beneficiaries understand their CVD risk and the benefits and drawbacks of different treatment options, then jointly deciding on a clinical approach to reduce risk that reflects the beneficiary's goals, values, and concerns; (2) annual risk reassessments (in person) to identify changes in each high-risk beneficiary's clinical risk and update his or her care plan; and (3) a minimum of two interactive follow-up contacts (any mode) each year to assist the beneficiary in making progress on the care plan.
- Collect and report clinical data to CMS via the Million Hearts Model Data Registry. The organizations will submit eligible beneficiaries' initial risk scores and supporting clinical data, as well as risk score updates over time (based on the longitudinal tool developed by Lloyd-Jones et al. [2017]) for high-risk beneficiaries.
- Participate in learning system activities, including webinars and videoconferences, designed to spread effective strategies for implementing the model, particularly through peer-to-peer learning.

CMS supports the intervention organizations with payments for risk stratification, cardiovascular management, and risk reduction. Participating organizations receive $10 for each eligible beneficiary they risk stratify. In the first model year, the cardiovascular management fees are a fixed $10 per beneficiary per month (PBPM) for each high-risk enrollees. In model year 2 and later, CMS is replacing the cardiovascular management fees with risk reduction payments that are scaled (up to a maximum of $10 PBPM) to the organization's performance in reducing 10-year predicted risk among their beneficiaries who were high-risk at initial enrollment. To support the model's evaluation, CMS is also paying control organizations to collect and report clinical data on their eligible Medicare FFS beneficiaries, but these organizations are not asked to calculate CVD risk scores or otherwise change their clinical care.

B. Objectives of the Million Hearts® CVD model impact evaluation

Our goal is to assess **whether the Million Hearts CVD model reduces the incidence of first-time heart attacks and strokes and CVD-related spending** for high- and medium-risk beneficiaries**.** The core design aims to estimate the impact of the Million Hearts CVD model as the regression-adjusted difference in outcomes among eligible medium and high-risk Medicare beneficiaries in the intervention and control groups. The evaluation will also test as secondary outcomes whether the Million Hearts model reduces overall Medicare spending (including with and without model payments), all-cause mortality, predicted CVD risk, CVD-related health care utilization (including hospitalizations, emergency department visits, and office visits), or use of CVD-related medications. We'll also examine intermediary measures of cardiovascular care, including changes in organization-level approaches to CVD care delivery.

[2] These criteria include being enrolled in Medicare Part A and B, not having end-stage renal disease, and not receiving hospice benefits.

II. DATA SOURCES

A. Primary data collection

1. Provider survey

We plan to survey intervention and control organizations and their providers to obtain self-reported measures of approaches to assess and mitigate CVD risk for their patients. The survey will cover topics such as proactive approaches to reducing CVD risk, risk stratification, individual risk modification, team-based care, population health management, quality and data reporting, and organization characteristics. Examples of these questions are shown in Table II.1.

Table II.1. Example questionnaire items to assess provider CVD care delivery

Assigning risk scores to eligible beneficiaries
What proportion of Medicare beneficiaries in your panel have you or your clinical team calculated a cardiovascular risk score for, using any risk calculator? (0; 1–24%; 25–49%; 50–74%; 75–100%; Don't know)
Thinking about the care you provided 2 years ago, what fraction of Medicare beneficiaries in your panel then did you or your clinical team calculate CVD risk scores for? (0- We did not calculate CVD risk scores 2 years ago; 1–24%; 25–49%; 50–74%; 75–100%; Don't know)
Provider awareness and use of risk scores
Are you, or is your clinical team, reviewing CVD risk scores for Medicare beneficiaries in your panel more consistently now than you were 2 years ago? (No change from before; Yes, somewhat more consistently; Yes, much more consistently; Don't know)
Is calculating CVD risk scores helping you identify Medicare beneficiaries in your panel as high risk who you did not previously recognize as being "high risk"? (Yes; No; Don't know)
Is calculating CVD risk scores helping you identify Medicare beneficiaries in your panel as medium risk who you did not previously recognize as being "medium risk"? (Yes; No; Don't know)
Once a risk score has been calculated, how often are CVD risk scores available when you meet with Medicare beneficiaries in your panel? (Always or almost always; Sometimes; Never; Don't know)
Notifying beneficiaries of risk scores
How are Medicare beneficiaries in your panel notified of their CVD risk score, if at all? SELECT ALL THAT APPLY ☐ In person at office visit, by provider ☐ In person at office visit, by other clinical staff ☐ Telephone call from provider ☐ Telephone call from other clinical staff ☐ Written communication (e.g., letter, email, patient portal)
Following up with high risk beneficiaries
Once you have identified Medicare beneficiaries as having high CVD risk, how often does your practice follow up with them through any mode (e.g., office visits, telephone calls, emails, or letters) to monitor plans to reduce risk? (Monthly or more often than monthly; Every 3 months, Every 6 months; Annually; As needed; Don't know)
Do you use any of the following resources to help ensure that your Medicare beneficiaries with high CVD risk are not lost to follow-up? SELECT ALL THAT APPLY ☐ Care managers ☐ Registries or tracking tools ☐ Automated scheduling of follow-up visits with a minimum frequency
New programs and services to treat CVD risk
In the past two years, has your practice added any new programs or services to address the following CVD risk factors in your practice's patient population? (Blood pressure control; Cholesterol management; Smoking cessation; Medication adherence; Changes in lifestyle, including weight loss and exercise)
General care delivery for CVD prevention
The cardiovascular preventive care our practice provides now is significantly different than the cardiovascular preventive care we provided before the CMS Million Hearts CVD Risk Reduction Model began in January 2017.
Participation in the CMS Million Hearts CVD Risk Reduction Model has prompted our practice to provide more systematically what is considered the current standard of care in this field.

Methodology. The survey features a mixed-mode survey administration approach. Respondents can complete the self-administered questionnaire either by web or with a paper version sent by mail. No financial incentive will be offered, because all organizations agreed to cooperate with the evaluation as a condition of participation.

B. Secondary data

1. Million Hearts Data Registry

We will acquire beneficiary- and provider-level data that Million Hearts CVD model participants submit to the Million Hearts Data Registry. We aim to use registry data to (1) define the beneficiary population for the impact analysis and assess baseline similarity of the intervention and control beneficiaries, and (2) define the secondary outcome of change in CVD risk score. We will acquire the following registry files through a shared folder CMS's subcontractor (Deloitte) on CMS's Chronic Conditions Warehouse (CCW).

1. **Demographics:** A person-level file with beneficiaries' age, sex, race, and other personal characteristics, including Health Insurance Claim Number, needed to link registry data to Medicare claims
2. **Visit:** A visit-level file with one row per model-related visit per beneficiary (expected maximum of one visit per person per year for high-risk beneficiaries)
3. **Clinical:** A file with clinical information, such as lab results or medication start dates, with one row per procedure or other clinical code
4. **Alignment:** A person-level file showing whether a beneficiary is model-eligible, is aligned with a given organization, and has complete baseline data
5. **Provider list:** A list of National Provider Identifiers (NPIs) associated with each model participant
6. **Organization list:** The list of participant organizations, with Taxpayer Identification Numbers (TINs) and organization contact information

We plan to use the alignment file to identify who enrolled in the model, and the demographic, visit and clinical file to identify when they enrolled, and their demographic and clinical characteristics at the time they enrolled. We will use the provider and practice file to identify who enrolled the beneficiary, and whether that participating provider and organization was in the intervention or control group.

2. Medicare FFS and Part D claims and enrollment data

a. Medicare data

We will use Medicare administrative data as the principal secondary data source for the impact analyses. The Medicare Enrollment Database (EDB) will provide information, by month, for beneficiaries enrolled in Medicare during the study period, including the parts of Medicare in which they were enrolled (Part A or B or a health maintenance organization—that is, Part C); whether Medicare was their primary payer of medical bills; and whether they were dually enrolled in Medicare and Medicaid. The EDB also will provide basic demographic data and the date, if applicable, a person died. Medicare FFS Parts A, B, and D claims will provide

information on service utilization, expenditures, medications used, and diagnostic history. We will acquire and process these data directly within the CCW VRDC to generate cost and utilization outcomes, as well as for linking the Million Hearts Data Registry with Part D event files to assess initiation and intensity of medication treatments for beneficiaries who have a Part D plan. This analysis will be limited to Medicare FFS beneficiaries with a Part D plan. Approximately 70 percent of Medicare beneficiaries are enrolled in a Part D plan.

We will use claims data with at least 90 days of runout, the standard for research purposes, and all data processes and programs will be subject to our rigorous quality assurance practices. We will use the VRDC's SAS GRID and multiprocessor environment to efficiently process multiple years of claims data, beginning with 2007 through the most currently available with 90 days of runout. We will request access to claims and enrollment data starting in 2007 (the earliest available on the VRDC) for use in determining whether a beneficiary has a history of acute myocardial infarction or stroke since 2007 or since they have been observable in Medicare claims data (if later than 2007).

3. Other CMS data

We will collect data from CMS, via the implementation contractor, on the actual amount CMS paid to each of the organizations (intervention and control) for participating in the model. These payments will both be important for assessing the impacts of the program on total Medicare spending. Specifically, we will compare any savings in Part A and B spending to the payments made to the participating organizations to determine whether the savings were enough to fully offset the cost of the program.

III. STUDY DESIGN

The core design for estimating impacts is a cluster randomized trial. CMS randomly assigned organizations (the clusters) to intervention and control groups. The intervention and control organizations were balanced on location (as defined by region), number of sites and practitioners, self-reported type of organization and estimated number of Medicare beneficiaries (NORC 2016). While the unit of random assignment was the organization, the unit of analysis for most study outcomes will be the beneficiary. That is, we will estimate impacts as the regression-adjusted differences in outcomes between intervention and control *beneficiaries*, using as the analysis sample all eligible Medicare FFS beneficiaries that the participating organizations reported in the Million Hearts Data Registry.

The time unit for most analyses will be the beneficiary-year. That is, we will follow each beneficiary for year 1, 2, 3 etc. after the date that they enrolled in the program (the enrollment date is reported in the registry). Because beneficiaries will enroll at different times, the follow-up years will cover different calendar periods for each beneficiary. We will use an intent-to-treat design, following beneficiaries for all months after they enter the Million Hearts CVD model, regardless of whether they continue to receive care from the intervention or control organizations over time. This will limit the possibility that differential attrition between the intervention and control groups will bias impact estimates.

We will consider two populations of beneficiaries separately: 1) high risk beneficiaries only and 2) high and medium-risk beneficiaries combined. CMS is paying organizations to conduct shared decision making and longitudinal cardiovascular care management for high-risk patients. So, the program could certainly be expected to improve outcomes for high-risk patients. However, the program could also reduce CVD events for medium risk patients because CMS is incentivizing organizations to risk stratify all of their patients.

We plan to present both Bayesian and frequentist impact estimates, including traditional *p*-values and confidence intervals, but also probabilistic statements more intuitive to policymakers, like "The program has a 70 percent probability of reducing the number of patients experiencing heart attacks or strokes by 10 percentage points or more over 10 years."

While the main impact analyses will estimate impacts on beneficiary outcomes, we will also estimate program impacts on the organizations' approaches to CVD care, as captured in the practice survey. For these outcomes, the unit of observation is the organization not the beneficiary. This means that the design simplifies to a standard (not clustered) randomized trial. Accordingly, we will estimate impacts as the regression-adjusted differences in outcomes between the intervention and control organizations. Further, for these outcomes, we will be measuring outcomes in calendar time rather than individual enrollment time.

IV. OUTCOMES

We plan to estimate impacts of the Million Hearts CVD model on two primary outcomes and a number of additional long-term and intermediate outcomes (Table IV.1). Most outcomes will be calculated separately among just high risk beneficiaries and among a combined sample of

medium and high risk beneficiaries. These outcomes will allow us to determine whether the expected causal chain of events occurred as planned and which events, if any, did not.

- **Primary outcomes**:composite incidence of first-time heart attack or stroke, CVD-related Medicare spending (for acute myocardial infarctions/stroke hospitalizations and related post-actue care and acute myocardial infarctions/stroke emergency department visits)
- **Other long-term outcomes**
  - **Health:** All-cause mortality, population-wide 10-year predicted CVD risk using the longitudinal risk assessment tool (Lloyd-Jones et al. 2017)
  - **Spending**: total Medicare spending, with and without program payments.
  - **Service utilization.** Hospitalizations for heart attack, stroke, and other cardiovascular disease; outpatient emergency department visits for heart attack, stroke, and other cardiovascular disease; number of office visits with a Million Hearts-participating provider
- **Intermediate outcomes**
  - **Medication use:** initiation or intensification of cholesterol or blood-pressure lowering medications. Data on medications will come from Part D but will be limited to the beneficiaries who are enrolled in Part D plans (nationally about 70 percent of Medicare beneficiaries are enrolled in Part D plans).
  - **Organization-level CVD care delivery:** Provider reported measures of organizational-level approaches to assess and mitigate CVD risk for their patients, such as the fraction of their Medicare beneficiaries for whom they assess CVD risk

Table IV.1. Outcome measures for impact evaluation, by domain, with data source and the preferred regression modeling approach

Domain		Potential measure	Data source	Regression model
Long-term outcomes				
Health	✓	First-time heart attack or stroke	Claims	Hazard model
		All-cause mortality	Claims	Hazard model
		Change in predicted CVD risk score[a]	Registry	Multi-level longitudinal regression
Spending	✓	CVD-related Medicare spending	Claims	Multi-level longitudinal regression
		Total Medicare spending without Million Hearts CVD model payments	Claims	Multi-level longitudinal regression
		Total Medicare spending with Million Hearts CVD model payments	Claims; CMS reported payments	Multi-level longitudinal regression
Service utilization		CVD-related hospitalizations	Claims	Multi-level longitudinal regression
		CVD-related emergency department visits	Claims	Multi-level longitudinal regression
		Million Hearts office visits	Claims	Multi-level longitudinal regression
Intermediate outcomes				
Medication use		Initiation or intensification of statins to lower cholesterol	Claims (Part D) linked to registry	Multi-level longitudinal regression
		Initiation or intensification of medications to lower blood pressure	Claims (Part D) linked to registry	Multi-level longitudinal regression
Clinical care		Clinical quality of care and processes, including: • Percentage of patients risk-stratified • Follow-up with patients who need CVD risk reduction • Adoption of certain CVD-reduction strategies, such as offering in-house smoking cessation therapy	Provider survey	Organization-level longitudinal regression

✓ = Primary outcome

[a]Risk scores will be calculated using the Million Hearts Model Longitudinal ASCVD Risk Assessment tool (Lloyd-Jones et al. 2017).

V. REGRESSION MODELING

A. Frequentist and Bayesian repeated outcome models

For most outcomes, we will have multiple observations per beneficiary (for example, 10-year predicted CVD risk in follow-up years 1, 2 and 3). For these outcomes, we will use a longitudinal, multilevel (hierarchical) mixed-effects regression model to compare regression-adjusted outcomes for the intervention and control beneficiaries (1) during each follow-up year and (2) averaged across all follow-up years. These regression models decompose the error term into the sum of a random error term for each organization, a random error term for each beneficiary, and a residual error term—thereby accounting for correlation of beneficiary outcomes across time for a given patient and across beneficiaries within an organization. When estimating the regression models, we will annualize the data (as needed) and weight observations to account for the share of the year for which each beneficiary's data are observed.

The specific regression model will be:

$$y_{ipt} = \sum_{\tau=1}^{T} 1(t=\tau) * \left(\alpha_\tau + \delta_\tau \, MH_p\right) + \beta \, x_i + \phi \, y_{i0} + \theta \, z_p + \gamma_p + \lambda_i + \varepsilon_{ipt}, \quad \text{Equation 1}$$

where y_{ipt} is the outcome measured for beneficiary i assigned to organization p in follow-up year t. In this equation, τ indexes the years (0 for baseline, 1 for follow-up year 1, and 2 for follow-up year 2, and so on), T is the maximum number of years available for the outcome measure at the time of the report, the function "$1(t=\tau)$" is an indicator function that is used to allow the regression coefficients to vary by year, and MH_p is a dummy variable that equals one for beneficiaries in intervention organizations and equals zero for beneficiaries in control organizations. Our hierarchical model decomposes the error term into the sum of a random error term for each organization, for each beneficiary, and a residual error term. In particular, γ_p is an organizationorganization-p-specific effect that accounts for clustering of beneficiaries within organizations, λ_i is a patient-i-specific effect that accounts for the correlation across repeated observations of beneficiary-i's outcomes, and ε_{ipt} is a time-varying residual error term.

The coefficient δ_t is our parameter of interest—it captures the impact of exposure to the program in year t. In the case of linear models, this coefficient can be directly interpreted as the program impact in year t—the regression-adjusted average difference in year t between intervention and control beneficiaries. We can also compute average impacts across the follow up years $\left(\bar{\delta} = \frac{1}{T}\sum_{\tau=1}^{T} \delta_\tau\right)$. For nonlinear models, we will calculate average marginal effects that express impacts in the same units as the outcome. (For example, for binary outcome measures we will use multilevel mixed-effects logistic regression models and express impacts as percentage point differences.)

The remaining covariates in Equation 1 are included to account for trends in the control group, improve the precision of the impact estimates, and net out effects of any observed differences in characteristics between the intervention and control groups that arose by chance despite randomization. The coefficient α_t captures the secular effect of patient-time in year t. For example, patient CVD risk increases as patients age, all else equal. The coefficients β, ϕ, and θ control for the effects of patient- and organization-level covariates measured at baseline

(x_i and z_p), respectively, and baseline outcomes y_{i0} when available. (Baseline outcomes are generally available for outcomes measured through claims and the registry, but not the survey.) The beneficiary covariates (x_i) will account for the date of enrollment, beneficiary demographics, Medicare and Medicaid enrollment status, beneficiary health status, non-modifiable CVD risk factors at baseline, and baseline outcomes measured in pre-enrollment claims data (when applicable). For registry-based outcome measures such as the CVD risk score, we could also account for the date within the follow-up year when the measurement occurred. Organization-level covariates (z_i) would include, organization characteristics (for example, organization size or-primary care versus specialty), geographic region, and participation in other CMS initiatives at baseline.

In the case of one observation per beneficiary ($T = 1$) as we will have in the beneficiary survey, our model reduces to the following random intercept model:

$$y_{ip} = \alpha + \delta\, MH_p + \beta\, x_i + \phi\, y_{i0} + \theta\, z_p + \gamma_p + \varepsilon_{ip}, \qquad \text{Equation 2}$$

where the terms are defined the same as in Equation 1 and "t" is suppressed for simplicity. Notice the beneficiary effect, λ_i, drops out of the model because the model contains just one observation per beneficiary. However, we continue to account for beneficiaries being clustered in organizations (γ_p).

For outcomes collected from the provider survey (that is, organizations' approaches to CVD care delivery), we will use a similar regression model, but the unit of analysis will be the organization, not the beneficiary. The regression models will include organization random effects to account for the correlation between repeated observations of a organization's outcomes. Unlike the patient-level analysis of claims data, this analysis will be subject to potential nonresponse bias (that is, due to nonresponse to the provider survey), so we will develop and use nonresponse weights to adjust for this possibility.

Specifically, we would use the following random intercept model:

$$y_{pt} = \sum_{\tau=1}^{T} 1(t = \tau) * \left(\alpha_\tau + \delta_\tau\, MH_p\right) + \beta\, x_p + \theta\, z_p + \gamma_p + \varepsilon_{pt}, \qquad \text{Equation 3}$$

where the terms are defined the same as in Equation 1, except that beneficiary-level covariate data would be aggregated to the organization level before being included in the model (x_p). If outcomes are observed multiple times per organization, the regression models for organization-level outcomes will include organization random effects (γ_p) to account for the correlation across repeated observations of a organization's outcomes.

We will present all key impact estimates in both Bayesian and frequentist frameworks. The model results will yield traditional *p*-values and 90-percent confidence intervals, using standard errors of the impact estimates that account for clustering. Further, because the outputs from these models closely approximate full Bayesian models (assuming flat priors), we can also make more intuitive probabilistic statements like, "There is a 75 percent chance that the intervention reduced CVD risk by 5 percentage points or more by the second year of follow-up."

B. Time-to-event analysis

For the primary outcome of heart attack and stroke and the secondary outcome of mortality, we will use hazard modeling to estimate impacts on the incidence of these events, as well as on the time to these events, given that the model could significantly delay these events even if—over a long time horizon—it does not prevent them altogether. Specifically, we will use a Cox proportional hazard model with "shared frailty." A shared-frailty model is the survival-model analog to regression models with random effects—in this case used to model correlations for beneficiaries within the same organization. Cox proportional hazard models are widely used in biostatistics to model impacts on event data. A major advantage of this model is that it uses data for all beneficiaries—even those who do not have data for the full test period because they enrolled in the Million Hearts CVD model later in the intervention period, or because they died before experiencing a heart attack or stroke. We will use the output of the model to plot the probability of having a first-time heart attack or stroke, by year of patient follow-up, for the intervention and control groups—with the differences in these probabilities reflecting estimated program impacts.

The Cox proportional hazards model can be expressed as:

$$h_{ip}(t) = h_0(t)\,\gamma_p \exp\left(\delta\, MH_p + \beta\, x_i + \theta\, z_p\right), \qquad \text{Equation 4}$$

where $h_{ip}(t)$ is the hazard for beneficiary i in organization p (that is, the estimated probability of the event occurring at time t), $h_0(t)$ is a baseline hazard (which does not need to be known for us to estimate the other model parameters), γ_p is the organization-level frailty (a latent random effect), and the remaining terms are defined the same as in Equation 1.

The coefficient δ captures the effect of the Million Hearts CVD model on the time-to-event, adjusted for other predictors in the model. After we fit the model, we will calculate the cumulative hazard as a function of time for the intervention and control groups (for example, the estimated percentage of beneficiaries with an heart attack or stroke within one year of enrollment) and the relative hazard (hazard ratio) for intervention versus control beneficiaries.

Cox proportional hazard models requires a proportional-hazards assumption, which means that the survival curves for the intervention and control groups have hazard functions that are proportional over time which may not hold. Thus, we will conduct two supplementary analyses to support the main modeling approach. First, we will formally test the proportional-hazards assumption is not violated. If the assumption is violated, we will fit a hazard model that allows the relative hazard functions for the intervention and control groups to differ over time. Second, we will use Equation 1 to estimate the effects of the Million Hearts CVD model on the proportion of beneficiaries with the event during a specified period or the average number of events, using the subset of beneficiaries who can be observed for the full period. For example, we will estimate effects on the proportion of beneficiaries who had a heart attack/stroke within two years of enrollment for the beneficiaries who enrolled early enough to be followed for two years in available claims data. While some beneficiaries will be dropped under this alternative modeling approach, we expect the vast majority of beneficiaries will be enrolled in the first year of the Million Hearts CVD model and, by extension, be included in the analysis.

C. Subgroup analyses

We will estimate the impacts of the Million Hearts CVD model for key subgroups of organizations and beneficiaries where we might expect impacts to differ. A primary subgroup analysis will be high-risk versus medium and high-risk beneficiaries. Further, we will test impacts for subgroups defined by the share of their total 10-year CVD risk that is due to modifiable (for example, blood pressure, cholesterol) vs. non-modifiable factors (for example, age, gender), with the expectation that model impacts will be largest for beneficiaries for whom most of their CVD risk is modifiable. Other key subgroups of interest are defined by beneficiaries' gender, age, or race; beneficiaries enrolling earlier versus later in the intervention period (because impacts might grow as organizations gain experience); characteristics of the organizations serving the beneficiaries such as size, specialty, or planned implementation strategy; and organizations near or far from the "ceiling" of optimal CVD performance at baseline. Traditional frequentist methods for subgroup analyses can be very imprecise, particularly when subgroups are small, and have a risk of identifying unrealistically large "effects" due to noise. To overcome this challenge, we will use Bayesian analysis to examine the variation in impacts across subgroups of interest in the evaluation. The Bayesian approach is ideally suited to this research question because it enhances statistical precision when estimating impacts in small subgroups while implicitly adjusting for multiple comparisons, obviating the need for post hoc corrections that sap statistical power.

Another important subgroup analysis will be to estimate impacts for beneficiaries enrolled by organizations of different sizes. The organizations vary substantially in their self-reported number of Medicare FFS beneficiaries, with some organizations seeing a very large (over 10,000) number of Medicare FFS beneficiaries. This skewness in organization size will mean that our estimates of the overall impact of the program will be driven mainly by the impacts within large organizations, which account for the bulk of the beneficiaries in our analysis population. We believe this is appropriate because larger organizations will, in fact, dominate the true impact of the Million Hearts CVD model on total incidence of heart attacks and strokes among enrolled beneficiaries. However, the overall estimate might mask impacts for smaller organizations that will receive smaller weight in the overall estimate. Therefore, we will obtain estimate impact estimates for subgroups of beneficiaries from smaller organizations, using the Bayesian approach. We may also test for interactions between organization size and impacts, which may improve statistical power to detect effects by using all observations to assess the relationship between size and impacts.

ADDENDUM TO STUDY PROTOCOL AND STATISTICAL ANALYSIS PLAN

MAY 2023

As noted in the online supplemental materials to the journal submission, we added (1) a secondary outcome of first-time CVD events and CVD deaths, and (2) analyses of effects on mortality by cause. We did not specify these in the initial trial protocol from February 2019 because both require data from the National Death Index (NDI) and, in 2019, we did not anticipate having the necessary linked NDI-Medicare data. Those linked NDI-Medicare data became available in 2022. We believe the additional outcome and analyses are valuable to the study because some fatal CVD events do not generate a Medicare claim. This was a particular concern during the COVID-19 pandemic, which was also unforeseen in early 2019.

To stay within the CMS budget for this study despite adding work, we did not conduct the Bayesian analysis noted in the initial trial protocol.

REFERENCES

Centers for Disease Control and Prevention (CDC). "CDC Grand Rounds: The Million Hearts Initiative." *Morbidity and Mortality Weekly Report*, vol. 61, no. 50, 2012, p. 1017.

Centers for Medicare & Medicaid Services Center for Medicare and Medicaid Innovation. "The Million Hearts® Cardiovascular Disease (CVD Risk Reduction Model Request for Applications." September 9, 2015.

Lloyd-Jones, D.M., M.D. Huffman, K.N. Karmali, D.M. Sanghavi, J.S. Wright, C. Pelser, M. Gulati, F.A. Masoudi, and D.C. Goff, Jr. "Estimating Longitudinal Risks and Benefits from Cardiovascular Preventive Therapies Among Medicare Patients: The Million Hearts Longitudinal ASCVD Risk Assessment Tool: A Special Report from the American Heart Association and American College of Cardiology." *Journal of the American College of Cardiology*, vol. 69, no. 12, 2017, pp. 1617–1636. doi:10.1016/j.jacc.2016.10.018.

NORC 2016. "Quality Control Results – CMS Million Hearts Randomization Results". Excel file submitted to CMS in 2016.

Sanghavi, Darshak M., and Patrick H. Conway. "Paying for Prevention: A Novel Test of Medicare Value-Based Payment for Cardiovascular Risk Reduction." *JAMA*, vol. 314, no. 2, 2015, pp. 123–124. doi:10.1001/jama.2015.668