Primary Outcome Measure

Title: Time from randomization to first participant-reported COVID-19 infection.
Description: Reported in a post-randomization survey.
Time frame: From date of randomization to the earliest of either the first participant-reported COVID-19 infection up to and including 365 days after randomization or right censoring up to 365 days after randomization. Censoring events include (1) the last date of a completed post-randomization survey if no post-randomization surveys report COVID-19 infection and the completion date of the last post-randomization survey was ≤365 days after randomization, (2) 365 days after randomization if no post-randomization surveys report COVID-19 infection up to and including 365 days after randomization and the completion date of the last post-randomization survey was >365 days after randomization, or (3) date of study withdrawal.

Primary analysis: We will use a multivariable Cox proportional hazards regression model to estimate an adjusted hazard ratio of the effect of moderate versus low dose vitamin D on the outcome. Covariates will include age, sex, race and ethnicity, baseline sun exposure, baseline insomnia severity, baseline COVID-19 exposure outside work, baseline employment, randomization date, and baseline geographic location. The selection and coding of these covariates mirror a prior NIH-funded trial of vitamin D and COVID-19 (NCT04868903). Coding of these covariates include age (continuous), sex (male versus female), race and ethnicity (non-Hispanic White versus non-White or Hispanic), sun exposure (average ≤30 minutes per day versus >30 minutes per day), insomnia severity (no or subthreshold insomnia versus moderate or severe insomnia), COVID-19 exposure outside work (no versus yes), and employment (no versus yes). Non-Hispanic White participants will be defined as those who report White race and non-Hispanic or Latino ethnicity. Non-White or Hispanic participants will be defined as those who report either non-White race or Hispanic or Latino ethnicity. Whereas recruitment in the NIH trial was limited to Chicagoland and stratified by recruitment site (UChicago Medicine and Rush University Medical Center), this study recruited participants across the United States, such that we will include baseline geographic location as a covariate, defined by baseline mailing address latitude <40.0 or ≥40.0 degrees. Furthermore, whereas the NIH trial recruited participants who concluded their participation prior to when Omicron arrived in Chicago about December 1, 2021, every participant in this study could be active in the study after Omicron arrived in the United States about December 1, 2021. Rather than control for randomization date with a linear time trend and an indicator variable for potential participation after Omicron arrived like the NIH trial, we will control for randomization date with a series of thirteen quarters defined by calendar quarter and year (Q1 2021 [December 2020 through March 2021], Q2 2021 [April 2021 through June 2021], Q3 2021 [July 2021 through September 2021], Q4 2021 [October 2021 through December 2021], Q1 2022 [January 2022 through March 2022], etc. up to Q1 2024 [January 2024 through April 2024]).

Secondary analysis of non-Hispanic White participants living north of 40.0 degrees: We will use a multivariable Cox proportional hazards regression model to estimate an adjusted hazard ratio of the effect of moderate versus low dose vitamin D on the outcome in participants who are non-Hispanic White and whose baseline mailing address latitude is ≥40.0 degrees. Covariates will mirror those described in the primary analysis, excluding race and ethnicity and baseline geographic location. We prespecify this secondary analysis prior to unblinding because

the NIH trial cited above found stronger evidence of increased time to first COVID-19 infection with vitamin D supplementation among non-Hispanic White participants, and participants in that trial were recruited in the Chicagoland area which has a latitude north of 40.0 degrees. The NIH trial also found stronger evidence of increased time to first COVID-19 infection with vitamin D supplementation among participants who reported an average of ≤30 minutes of sun per day.

Additional Analyses

Time-varying analysis: We will use a time-varying, multivariable Cox proportional hazards regression model to estimate an adjusted hazard ratio of the effect of moderate versus low dose vitamin D in each calendar quarter (Q1: January through March, Q2: April through June, Q3: July through September, Q4: October through December) on the outcome in participants whose baseline mailing address latitude is ≥40.0 degrees. Covariates will mirror those described in the primary analysis, with the addition of four calendar quarter indicator variables as time-varying covariates. Furthermore, instead of using a single indicator variable for moderate dose vitamin D, we will include four indicator variables for treatment with moderate dose interacted with the time-varying calendar quarter. The reference group for the uninteracted and interacted time-varying calendar quarter indicator variables will be Q1 for low dose vitamin D.

Treatment heterogeneity analysis: We will use a multivariable Cox proportional hazards regression model to estimate an adjusted hazard ratio of the effect of moderate versus low dose vitamin D on the outcome in every subgroup of participants defined by the covariates in the primary analysis, except for the thirteen quarters defined by calendar quarter and year. Age will be dichotomized as <45 or ≥45 years. We will also analyze subgroups of Blacks or African-Americans, non-Whites or Hispanics with baseline mailing address latitude ≥40.0 degrees, non-Whites or Hispanics with baseline mailing address latitude <40.0 degrees, non-Hispanic Whites with baseline mailing address latitude <40.0 degrees, participants unimmunized for COVID-19 at baseline, and participants fully immunized for COVID-19 at baseline. Full immunization refers to one vaccine dose from Johnson & Johnson or two vaccine doses from other manufacturers.

Additional outcomes: We will analyze COVID-19 infection severity for moderate versus low dose vitamin D as measured by rates of symptoms using a modified Beat COVID-19 instrument (NCT04337762), hospital admission, hospitalization with ICU admission, hospital length of stay, whether ventilator, high-flow nasal oxygen, or helmet-based ventilation is used, and death.

Other Statistical Considerations

Proportional Hazards Assumption: In the primary analysis, secondary analysis of non-Hispanic White participants living north of 40.0 degrees, time-varying analysis, and treatment heterogeniety analysis, the proportional hazards assumption will be assessed using the *estat phtest* command in Stata with rejection of proportion hazards defined by $p<0.05$ for any variable or the global test, or by visual assessment of Schoenfeld residuals. Continuous variables which fail proportional hazards will be categorized. For any categorical variable or categorized continuous variable which fail proportional hazards, we will perform Cox regression which stratifies based on that variable. If the moderate dose vitamin D variable fails the proportional hazards assumption, we will estimate non-proportional hazards models.

Missing Data: Assuming missingness at random, we will impute missing values of covariates with multiple imputation by chained equations with 20 imputations. Covariates in the imputation will include the covariates in the primary analysis excluding moderate dose vitamin D. The imputation for the secondary analysis of non-Hispanic White participants living north of 40.0 degrees and the treatment heterogeneity analysis will be performed among the denominator of all participants in the primary analysis.

Statistical Significance and Multiple Testing: Statistical significance will be determined by $p<0.05$ for the primary analysis and the secondary analysis of non-Hispanic White participants living north of 40.0 degrees with no adjustment for multiple comparisons.

Software: Data analysis will be executed in Stata 18.0.