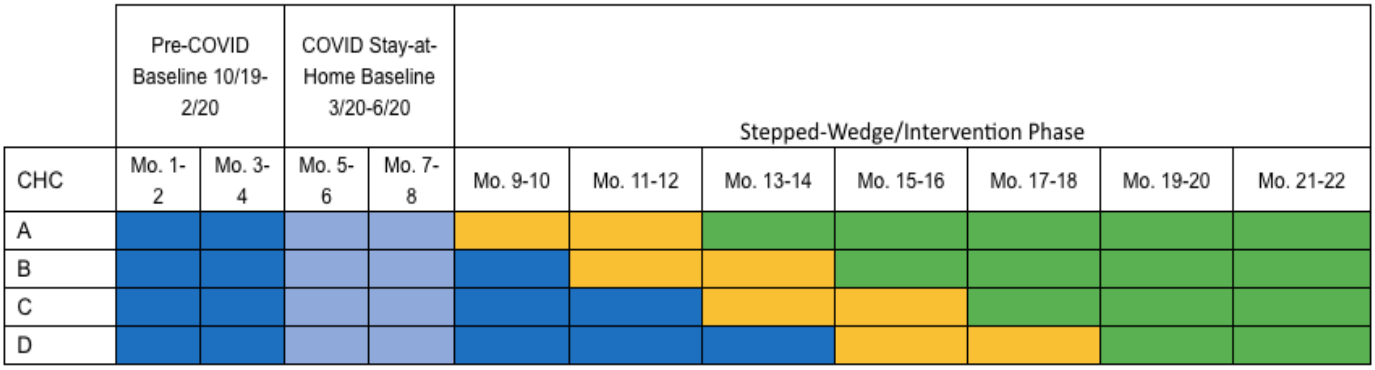
Paired Promotion of Colorectal Cancer and Social Determinants of Health Screening
NCT04585919
06/09/2021

Statistical Design and Power

Study Design: We have chosen a cross-sectional cluster stepped wedge study design. This pragmatic design involves a sequential, randomized roll-out of the intervention across the four participating clinical sites or "clusters" over 8 week intervals or "steps." The figure below details our approach: 8 months of historical controls at all of the sites (blue pre-COVID and light blue is COVID stay at home) sequential roll out of the base intervention every 8 weeks (yellow), followed by a period of planned adaptation (green).



Outcomes: The primary outcomes in our analysis will be: (1) return of FIT kit within 3 months; (2) completion of social needs screening; (3) referral of patients with social needs to appropriate resources.

Covariates: Age, race/ethnicity, primary language, household income (% of Federal Poverty Level, FPL), insurance, comorbid conditions, number of CHC visits in prior year, and date of prior CRC screening completion.

Analysis: To estimate the intervention effect, we will fit generalized linear mixed effects models (GLMM) controlling for secular trend. The basic GLMM approach to estimate intervention effects in stepped wedge studies includes an indicator for intervention and a linear or categorical input for time-period as fixed effects to control for secular trend. Given the possibility of intermittent social distancing, we will apply modifications to the basic model. First, we will examine non-linear secular trends by including splines for time-trend, which is needed if intermittent social distancing occurs and it affects outcome trends. Second, we will consider time-trend as a random effect, which allows the secular trend to vary across the CHCs. This may be needed if mandated social distancing occurs by region and differentially affects the CHCs' secular trend. Third, we will consider interactions for time-trend and the intervention term (to allow time trend to vary pre- and post-intervention) and patient characteristics (for differential time trends by characteristics). This may be needed if subgroups (e.g. those over 60) are asked to intermittently social distance. As we cannot predict the social distancing recommendations moving forward, we describe in general terms how we will modify the basic modeling approach. Our strategy addresses the potential concerns with the basic GLMM model for stepped wedge designs in the present health climate [8]. Our modeling strategy will be informed by the judgment of the substantive experts on the investigative team with input from the CHCs and comparing models statistically using the likelihood ratio tests. The modeling approach will be used for each of the three primary outcomes. We will consider intervention effects statistically significant if we observe a p-value below 0.017 which correspond to $\alpha = 0.05$ threshold with a Šidák correction for three comparisons. If specific sub-groups of the patient population are particularly affected by social distancing, we will explore differential treatment effects by including interaction of intervention term and patient characteristic to the final GLMM model, as described.

Statistical power and sample size: We estimate that there will be 46,650 adults who need CRC screening across all of the study sites for the supplement project period and one pre-supplement period control step. Because of the nature of the CHC patient population, all patients' situations will likely have changed and thus should receive a social needs screening. Further, we estimate that virtually all patients will have some increased social needs. To estimate power for our sample size we conducted a simulation study under three scenarios for secular trend – (1) linear small positive or flat secular trend and an intervention effect of 5% increase (3) a moderate cubic secular trend and an intervention effect of 5% increase, and (3) strong cubic

secular trend and an intervention effect of 7.5% increase. For each of these we assumed a conditional ICC of 0.10 and a baseline outcome of 50%. Each scenario incorporated the number of steps in our design, was simulated 1,000 times and modeled using the respective modifications to the basic model described above (see “Analysis”). Based on the sample size and our assumptions, we have greater than 90% power for scenarios (1) and (2), 87% power for scenario (3) at a Šidák corrected family-wise error rate of 5%. As such it is clear we have adequate power to detect a clinical meaningful effect size.