Cover Page

Document:	Statistical Analysis Plan
Document Date:	2020-12-07

Project Title:	COVID-19: Healthy Oregon (Oregon Saludable): Together We Can (Juntos Podemos)
Identifiers:	NCT04793464
Unique Protocol ID:	10032020.002
Secondary ID:	P50DA048756-02S2

4.3 Statistical Design and Power

4.3.a Optimization and Adaptive Research Design

The goal of this research approach is to test optimal strategies to increase SARS-CoV-2 testing in Latinx populations. The primary approach will be a sequential multiple armed randomized trial (SMART). The proposed SMART design will evaluate adaptive intervention strategies (AIS) which vary the testing delivery format and sequence of testing strategies tailored to Latinx communities. Our research will identify an optimal AIS at the site and individual level as well as explore downstream site and individual differences moderators for further personalization. Through the identification of an optimal AIS, our research will provide public health officials and health practitioners with data on specifically what is the most effective and efficient sequence and modal delivery of testing for underserved Latinx communities. Adaptive intervention strategies involve decision rules that specify how the type or intensity of a testing approach should change site response rates during the progress of implementation monitoring.[33,68]

Moreover, our research design is a hybrid investigation combining effectiveness evaluation and implementation research to further understand what elements of implementation lead to sustainability of the optimal testing strategies.[28] Because each AIS is designed to be a deliverable product for implementation, intervention approaches do not separately involve randomization. Rather, in the implementation of a SMART designed to evaluate these AISs, sites follow a two-stage randomization procedure in order to be assigned to each AIS. All sites are initially randomized to one of the two first-stage[33,34] interventions. Sites are then re-randomized to one of the two second-stage intervention strategies. This randomization procedure allows an evaluation of the relative effectiveness of each overall AIS as well the ability to explore other questions such as which first-stage intervention is the most effective and which second-stage intervention format is the most effective for slow-responding sites. Unique analytic strategies are used to maximize power to address these questions within SMART designs (see Statistical Analysis Plan below).

Within the SMART, a measure of initial response is assessed following implementation of the primary outreach approaches in order to evaluate each site's need for further adaptation. In this way, testing is tailored to meet the needs of communities and their Latinx individuals and families by providing enhanced testing strategies to slow-responding sites. This measure of initial response is referred to as a primary tailoring variable. In the proposed study, the primary tailoring variable is whether a site is operating at a threshold of 80% testing capacity, based on the frequency and per capita testing rates and presumption of 50 tests per testing event. The AIS each represent distinct plans of testing outreach approaches. Within the SMART, it is then possible to evaluate which adaptive sequence is the most effective in impacting Latinx testing rates and secondary outcomes such as health maintenance behaviors.

Specifically, the proposed RADx SMART includes 4 embedded adaptive intervention strategies (AIS). Sites and individuals in the proposed Latinx sampling frame will receive one of two first-stage implementation outreach strategies: Basic Outreach approaches or Basic Outreach plus Promotoras Comunitarias. Primary outcomes will be site level testing rates and individual level COVID-19-related health behaviors. After the initial 4 weeks (2 testing events), sites will be re-randomized in the second-stage randomization to either stay the course stationary events or to a mobilization intervention. Thus, for the primary criterion outcome of testing capacity, the SMART design includes 4 embedded site level adaptive intervention strategies (AIS): (1) starting with culturally adapted Basic Outreach, and staying the course; (2) starting with Basic Outreach followed by mobilization intervention; (3) starting with Basic Outreach plus Promotoras Comunitarias and staying the course; and (4) starting with Basic Outreach plus Promotoras Comunitarias plus mobilization intervention. In addition, as a booster to the Promotoras Comunitarias intervention, for health behavior outcomes, individuals tested at Basic Outreach sites will receive pamphlet health promotion pamphlets, and those tested at Promotoras Comunitarias sites, will receive a health pamphlet plus a Promotoras de Salud intervention.

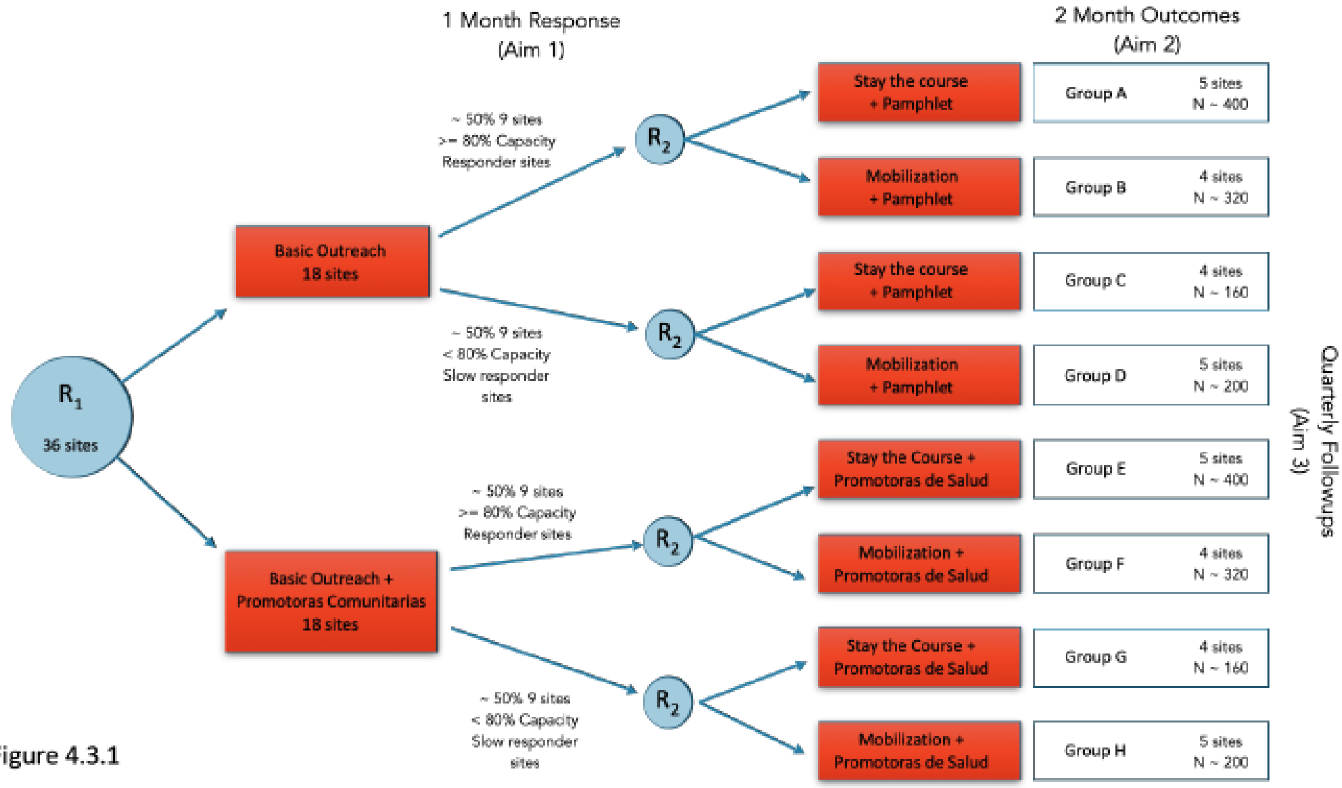

Figure 4.3.1

The use of a primary tailoring variable representing initial response allows the AIS to be responsive to individual Latinx family needs for effective testing. Secondary tailoring variables represent an additional approach to more deeply personalize AIS. Secondary tailoring variables include site or individual characteristics (i.e., moderators) and are used to examine the overall adaptive intervention strategies most likely to be effective from individual differences perspectives. Potential secondary tailoring variables are measured at baseline and evaluated following completion of the study for their ability to predict differential response to each AIS. For example, the SMART could reveal that sites in more socially disadvantaged communities or individuals with greater perceived stigma may benefit most from AIS 4 involving two Promotoras interventions allowing for maximum engagement outreach. Secondary tailoring variables, or identified moderators can be utilized for further adaptation in future implementation roll out. The study design is depicted in Figure 4.3.1 above.

4.3.b Embedded AIS Research Questions and Aims

The specific research questions and aims are the following. Question 1a: Is Basic Outreach plus Promotoras Comunitarias more effective for maximizing testing capacity than Basic Outreach? Question 1b: For individuals tested, is Pamphlet plus Promotoras de Salud more effective than Pamphlet only in supporting positive health behaviors for individuals tested? [(A+B+C+D) vs (E+F+G+H)]. Research Aim Question 2: Among slow responding sites, does the mobilization intervention yield an increase in individual positive health behaviors and in site testing capacity? [(D+H) vs (C+G)].

Each of these AIS research questions is depicted in Figures 4.3.2 through 4.3.3:

Question 1a: Is Basic Outreach plus Promotoras Comunitarias more effective for testing capacity than Basic Outreach alone, for Latinx communities?

Question 1b: For individuals tested, is Pamphlet plus Promotoras de Salud more effective than Pamphlet only in supporting positive health behaviors for individuals tested, and increasing site testing rates?

Aim 1a: Is Basic Outreach plus Promotoras Comunitarias more effective for testing capacity than Basic Outreach alone?
Aim 1b: For individuals tested, is Pamphlet plus Promotoras de Salud more effective than Pamphet only in supporting positive health behaviors for individuals tested, and increasing testing rates?

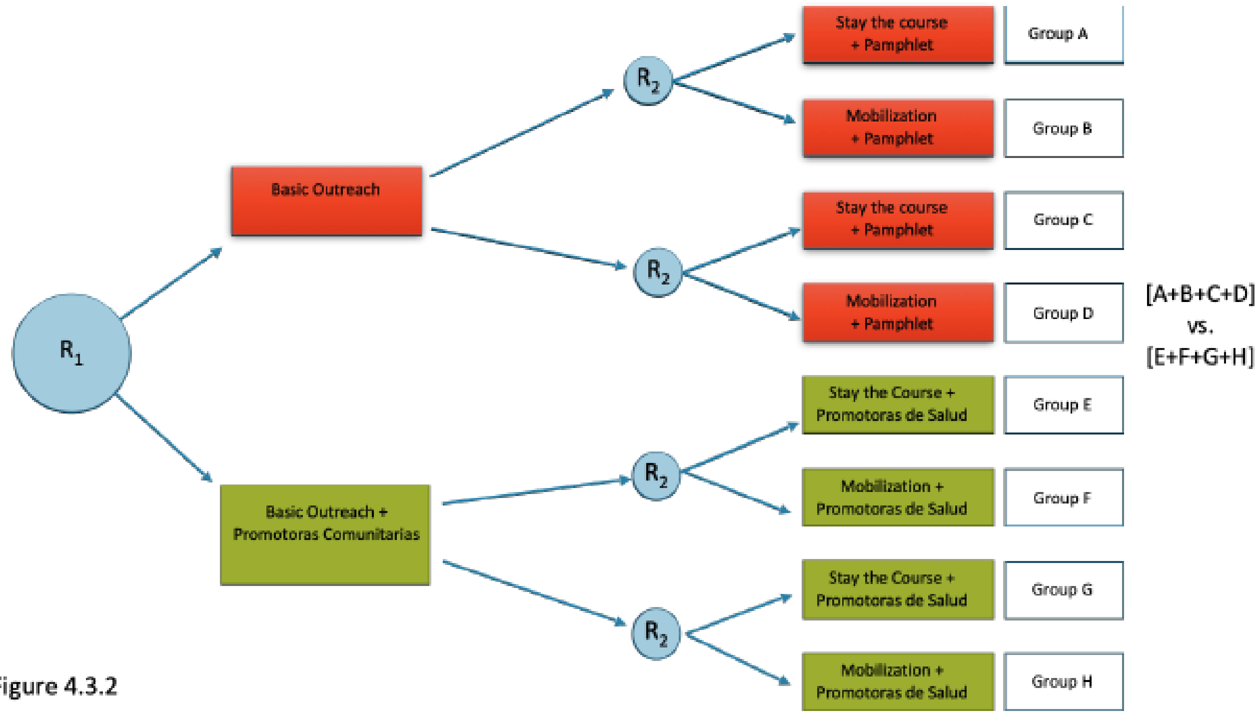

Figure 4.3.2

Research Aim Question 2: Among slow responding sites, does the mobilization intervention yield an increase in site testing capacity?

Aim 2: Among slow responding sites, does the mobilization intervention yield increase in site testing capacity?

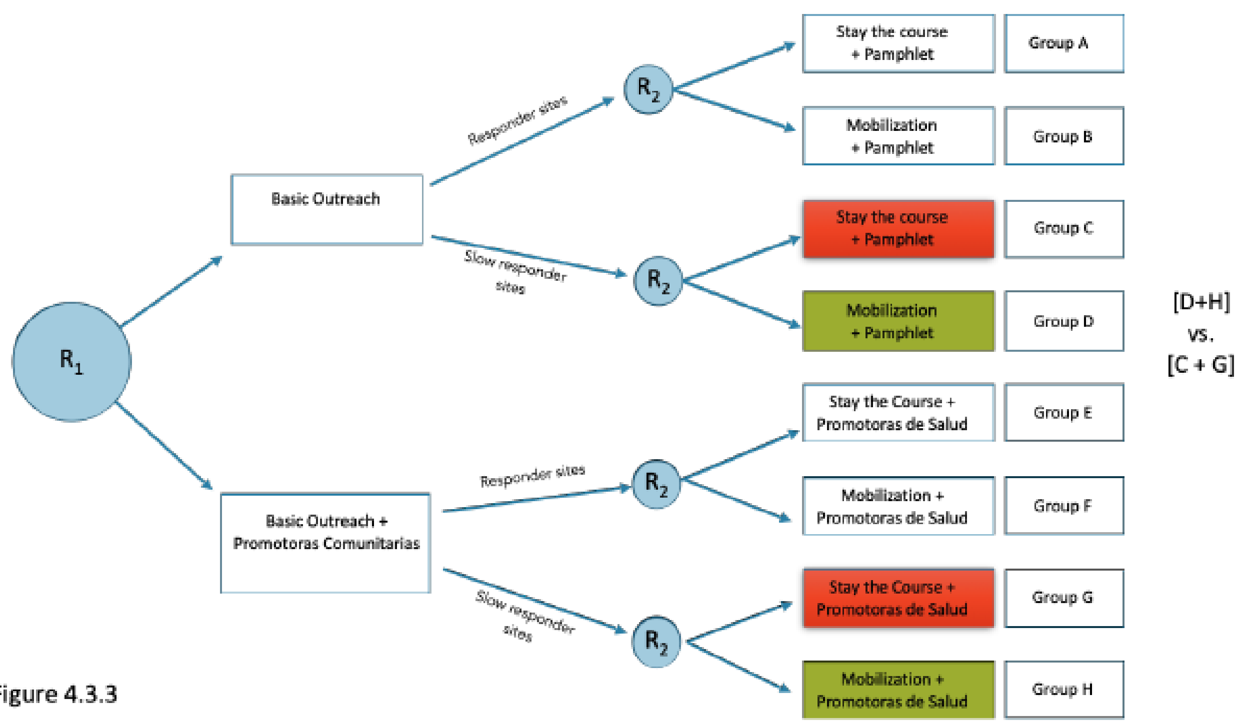

Figure 4.3.3

In addition to research questions 1 and 2, the SMART design also allows for an omnibus comparison; after starting with Basic Outreach or Basic Outreach + Promotoras Comunitarias, we can examine which of the embedded adaptive intervention strategies is more effective overall for testing capacity and health behavior maintenance [A vs B vs C vs D vs E vs F vs G vs H]. Although not a specific aim of the proposed study, agencies and stakeholders may be particularly interested in this question and our study design will allow us to test this, focusing on the combined R1 and R2 strategies.

4.3.c Recruitment of 36 Communities in 6 Counties into the Testing Program

This testing program will build on existing partnerships with 6 county public health authorities in Oregon (see letters of support), that are either designated as a rural county by the Office of Management and Budget (Lincoln, Douglas, Josephine) or have a significant rural population (Lane, Marion, Benton). We will establish contracts with each county to provide them with staffing resources to support additional testing in Latinx communities. In addition, we will work with each county public health agency to develop a County Collaborative that includes key community stakeholders in the Latinx community, including those representing Latinx serving organizations such as non-profits and churches. The focus of the County Collaboratives is to engage the implementation process described in the implementation strategy. Leveraging the county collaboratives' knowledge of county communities, 6 communities (identified by zip code or zip code cluster) within each county will be selected for site implementation. Given that proximity is a critical component for reducing barriers to access in rural communities, and given the composition of the 6 counties, most of the 36 communities who participate in the testing program will be rural. We will employ a *priori* stratification by county and rurality to balance potential confounders, to reflect the hierarchical structure of the design in the analytic plan. Specifically, for the first randomization, within each county, 3 communities will be randomized to the treatment condition and 3 to the control condition (total of 18 per condition, 36 total).

4.3.d Data Analysis for Aims 1 and 2

For the primary evaluation aims for determining the optimal strategy for testing Latinx communities and the four adaptive intervention strategies, we will employ standard normal theory generalized linear modeling analysis of covariance fractional factorial analyses using pre-post assessment scores for site and individual level outcomes. Research Question 1 will focus on the combined effects of A1 randomization groups [(A+B+C+D) vs (E+F+G+H)]. Research Question 2 will focus on which A2 strategy works best for slow-responding sites [(D+H) vs (C+G)]. To address the non-independence of subjects in the embedded adaptive intervention designs and crossed factors, we will estimate linear mixed models with weighted least squares[69] using the design effects coding shown in Table 4.3.1.

The general equation to estimate the optimal strategy is given by

$$E[Y|A_1, A_2, O_1] = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \beta_3 A_1 A_2 + \gamma^T O_1$$

where *E* is the expected value of the outcome, A1 is the initial treatment randomization, A2 is the second stage treatment randomization and γT are the estimates for covariate O1, the pre-test covariate. For binary and count data we will employ generalized linear modeling and logit, probit, and poisson link functions for criterion outcomes such as frequency of tests, re-tests, and discrete event behaviors.

4.3.e Research Aim 3 Question: Characteristics of Sustainable Testing Sites and Sustained use of Effective Approaches

Qualitative thematic analyses will be used to code and segregate data derived from key stakeholder interviews and meeting notes from quarterly reports to categorize, synthesize, and interpret information. Responses will be categorized and coded for engagement themes (acceptability, penetration, adoption, fidelity) following a multi-step qualitative analysis process.[70] We will independently conduct first-level coding to establish broad engagement themes. Next, we will independently conduct the second level of coding, meeting several times throughout the process to discuss areas of disagreement until consensus is reached. At the conclusion of coding we will have established ratings for level achieved for each engagement theme. Next, to the extent there is sufficient variability in the engagement ratings, we will examine correlations between engagement ratings and the number of sites within a community that have sustained testing (e.g., 80% or greater capacity) and use of efficacious interventions (e.g., number of times data-based decisions made). We acknowledge power to detect significant meaningful correlations is low given this is a community-based research question.

Thus, within each community, we will also rank order sites by outcomes and examine mean engagement rating scores which will provide additional evidence as to which engagement factors are associated with successful sustainability.

4.3.f Power

Power for detecting the optimal AIS was estimated using the SMARTAR package 1.0.0 in R R[71]. We chose a power convention of .80 with a two-sided alpha level of .05 to reduce the probability of obtaining a false positive (Type-I error). For research question 1, balanced response rates of 50% of sites having adequate response and 50% slow responding, 36 sites for a continuous outcome (e.g., rates) can minimally detect a small difference of .27 among the combined group comparisons [(A+B+C+D) vs (E+F+G+H)], and power to minimally detect a medium effect of .54 (n=18 sites) [(D+H) vs (C+G)]. For the global test of the optimal treatment strategy (n=36 sites), the study is powered to minimally detect an effect of .40, a moderate effect. Power for detecting the optimal AIS was estimated using the SMARTAR package 1.0.0 in R[71].

We must also consider response scenarios that include either more slow-responding or even less slow responding sites before the tailoring intervention (i.e., 80% of the randomly allocated sites responded at capacity and 20% were below threshold; or alternatively, only 20% responded and 80% were below capacity initially). In each of these "unbalanced" response scenarios, power for research question 1 is the same as the balanced design above. Although we do not anticipate such a large range of slow responding sites based on our preliminary testing; for the slow-responding sites only, a bracket range of 80/20 response ratio (resulting in 8 slow responding sites [(D+H) vs (C+G)]) to a high of 20/80 ratio (resulting in 28 slow responding sites [(D+H) vs (C+G)]. In these scenarios, for research question 2, the study is powered .80 to minimally detect a difference of 1.3 (a large effect) or to minimally detect a moderate effect of .35 (a moderate effect), respectively, for 8 to 28 total slow responding sites for research question 2.

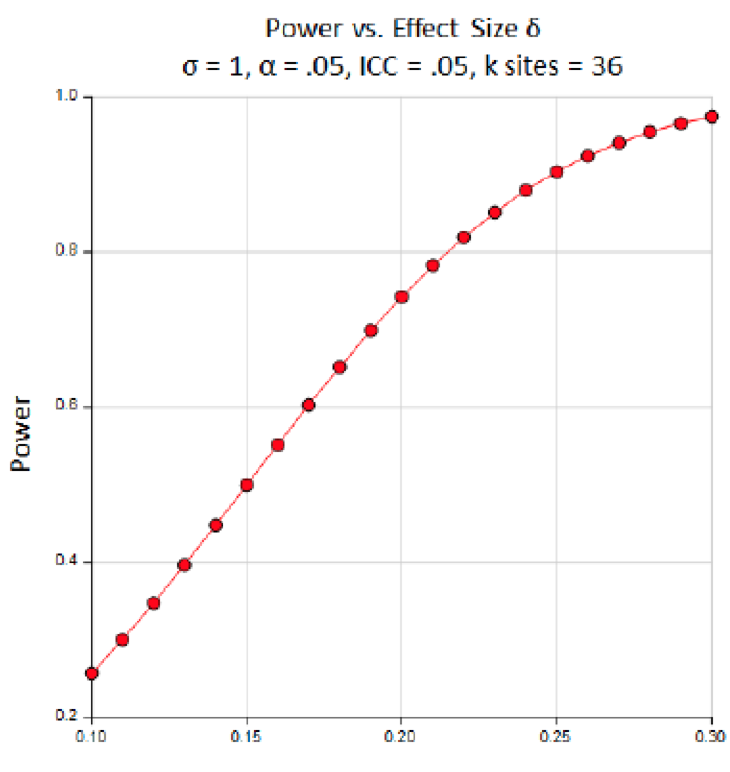

Figure 4.3.f Power Curve.

Finally, when considering power for individual level data, a mixed model with two levels (individuals nested within sites) we must consider full capacity scenarios where implementation is realized at full capacity (roughly 3600 tests) to scenarios that obtain a smaller number of individual tests based on response patterns by site. In this case, that is the balanced scenario shown in Figure 4.3.1 given half of the sites are slow responding. That scenario projects a maximum of 2520 level-1 testing units sampled from 36 sites and achieves .80 power to detect a difference of .22 (a small effect) among the group conditions for randomization 1 conditions A1, randomization 2 conditions A2, and their crossed A1A2 interaction, and an intraclass correlation of .05 for the test of comparative intervention strategies from the primary equation shown above. Thus, the study is adequately powered for all multilevel data scenarios for minimum response to maximum response to detect small effect sizes. Power for individual comparisons among the 4 embedded AISs is shown in Figure 4.3.f.