NCT05934591 Unique Protocol ID: UIUC IRB 23828
Improving Attendance in Community Wise
Statistical Analysis Plan 6/1/2025

Statistical Analysis Plan

Variables: marital status, education, number of days paid in the past 30 days, job search status, and employment status. For modeling, potential covariates were selected based on their association (p value <0.2) with one or more of the experimental conditions (surveillance, incentivizing intervention attendance, group type) or attendance rate. We allowed for a larger p value cut off given the small sample size. Associations between continuous measures were examined using Spearman correlation coefficients. Difference in continuous variables between groups were assessed using the Wilcoxon 2-sample test. Differences in the proportion of categorical variables between groups were assessed using the Chi-Square or Fisher's Exact test.

To test the objective 'What combination of interventions was associated with higher attendance?', we used a logistic regression model to test the association of group type (open vs closed), incentivizing intervention attendance (incentive vs not) and surveillance (under surveillance vs not) with the outcome of attending 50% of sessions or more (vs not). The logistic regression model included main effects, as well as all 2-way interactions and a 3-way interaction using multiplicative terms. The model was adjusted for job search status, sex, race and ethnicity, marital status income, number of days paid in the past 30 days, and employment status. Parameter estimates (coefficients) and their standard errors were reported. Since the outcome for logistic regression model is the log odds of attending 50% or more sessions, parameter estimates (regression coefficients) represent the change in the odds of the outcome on the log scale for a one-unit change in the independent variable (or at one level compared to its referent if continuous). Thus, parameter estimates were transformed by taking the exponent to facilitate interpretation as odds ratio. Exponentiated coefficients are interpreted as the difference in the likelihood (odds) of achieving the outcome between those who received an intervention component compared to those who did not receive it, collapsing (averaging) over levels of the other components and controlling for covariates. To calculate percent change, exponentiated coefficients are compared to the value of one (one means no difference). A value that is greater than one or less than one, indicates that the outcome is more likely or less likely to occur, respectively. Thus, we subtract one from the exponentiated regression coefficient to obtain the magnitude of change in the outcome, and when multiplied by 100 it is a percent change in the outcome.

To test the objective 'Does higher attendance result in less ASM use?' we employed a generalized multiple regression model testing the association of attendance rates (continuous variable), incentivizing intervention attendance (incentive vs not), surveillance (under surveillance vs not) and their 2-way interaction with the outcome of percent of ASM at follow up, adjusting for demographic characteristics and percent of ASM at baseline. We analyzed the intervention components using effect coding (1 vs -

1). Thus, the parameter estimates (regression coefficients) that included intervention components were multiplied by a scaling constant of 2 to obtain the main and interaction effects as per guidelines for using effect coding in multiple regression. We did not include group type in this model because group type did not meet our selection criteria for inclusion in the optimized intervention. Due to the skewness of the ASM variable, we assumed a negative binomial distribution, thus we calculated the exponent of these scaled parameter estimates to facilitate interpretation as the percentage of ASM use. For interpretation, exponentiated coefficients are compared to the value of one to estimate a percent change, as was explained for the logistic regression model.

For all models, fit was assessed using standard criteria including the Akaike Information Criterion (AIC), Pearson Chi-Square, and Wald test. Analyses were conducted using SAS version 9.4 (SAS Institute, Cary, NC), using two-sided tests and a significance level of 0.05.

Using the causal mediation approach, we tested the mediation of incentivizing intervention attendance on the relationship between attendance with the outcome of percent of ASM at follow up, adjusting for demographic characteristics and mean percent of ASM use at baseline. Since no mediation effect was noted, no results are shown.