PRS Cover Page for Statistical Analysis Plan

Title: Parent Feedback Intervention Targeting Student Transitions and Alcohol Related Trajectories (+) Efficacy Study (FITSTART+)

Date: April 26, 2021

ClinicalTrials.gov Identifier: NCT04549454

Statistical Analysis Plan

Examines the efficacy of FITSTART+. Main effects of the intervention will be examined with hierarchical generalized linear models (HGLM) using SAS. HGLM permits the simultaneous examination of relationships among variables at different levels, for example variables measured over assessment periods assumed to fluctuate within an individual (e.g., student drinking behavior) with variables considered to have stable characteristics (i.e., parent intervention condition). Using this framework, time will be specified as a Level 1 varying predictor nested within individuals (Level 2). The following model will be used to determine main effects of the intervention conditions compared to control:

Level 1: $DV_{ti} = \pi_{0i} + \pi_{1i}(T1)_{ti} + \pi_{2i}(T2)_{ti} + \varepsilon_{ti}$
Level 2: $\pi_{0i} = \beta_{00} + \beta_{01}(Tx)_i + r_{00i}$
$\pi_{1i} = \beta_{10} + \beta_{11}(Tx)$ t indexes repeated measure $\pi_{2i} = \beta_{20} + \beta_{21}(Tx)$ i indexes individuals

Respective model DVs will include total drinks per week from the DDQ, consequence composite score, and number of HED episodes during the past month. T1 and T2 will be dummy coded time variables (T1 = student baseline [coded 0] to 1 month into college follow-up [coded 1], T2 = student baseline [coded 0] to 6 months into college follow-up [coded 1]). Tx will be a categorical predictor comparing the intervention condition [coded 1] to control [coded 0]. With this model, intercept treatment differences will represent treatment differences at baseline (e.g., group differences in student drinking at baseline) and slope differences will represent changes over time (e.g., did students of intervention parents reduce their drinking between baseline and follow-up assessments more than students of control parents?). The intercept includes a random effect, which will model the subject-specific heterogeneity in the outcome and control for correlated data due to individuals. Pairwise comparisons (post hoc Tukey's) were conducted examine mean differences between conditions at each time point.

Missing data. Attrition bias will be tested by evaluating whether baseline outcomes differ between student and parent dropouts versus completers. HGLM is optimal in that it allows for unbalanced data, giving heavier weights to participants with more non-missing values. HGLM also assumes that missing values are missing at random as opposed to non-ignorable. The pattern-mixture approach to non-ignorable missing data will be used to assess the sensitivity of findings to the presence of missing data. If the pattern of missingness appears to pose a threat to the validity of the experimental manipulation (i.e., student attrition is related to parent random assignment, or the predictors of attrition are different for students of intervention vs control parents), more powerful missing techniques will be used, such as multiple imputation.

Additional Model Considerations. Student outcomes (drinks per week, HED frequency, and negative consequences) are count variables, which can have skewed distributions when the overall mean is low. Alternative models could also specify a Poisson or Gaussian distribution for the Level 1 errors. We will consider whether the

Level 1 components are more appropriately modeled as Gaussian or Poisson. Further, in spite of randomization at the parent level, we will conduct baseline equivalence tests between the FITSTART+ student participants to establish successful parent randomization effectiveness, including pre-treatment alcohol use, negative consequences, and demographic variables. Any student variables that are associated with treatment assignment, in spite of parent randomization, will be included as covariates.

Power. *We have more conservatively aimed for sufficient statistical power to detect small effects on our primary student drinking outcomes (i.e., HED frequency, drinks per week, negative consequences).* Very low student attrition (13%) was achieved in the original FITSTART study across 10 months and three assessments. As identical procedures and similar assessments will be used in the current study, we expect to retain over 85% of students. Power analyses were calculated using the Optimal Design program for longitudinal HLM specifying two-study conditions and randomization at the person-level. Power curves with 20% attrition were on plotted for sample size per treatment arm *at two effect sizes: d=.20 (small effect)* and *d*=.40 (medium effect). *A sample size of 300 per arm was found to yield excellent power for detecting both effect sizes.*