Study Title: Computerized Tool for Preventing Prenatal Drinking

NCT Number: NCT02337361

Date of Document: May 1, 2015

Approved: June 15, 2015

STATISTICAL ANALYSIS PLAN

Statistical Analysis Plan

Preliminary analysis will a) examine distributions of study variables, transform or categorize non-normally distributed variables, and handle outliers as recommended (Andrews et al., 1972); b) conduct attrition analyses to assess potential bias due to participant loss at follow-up (Gerstein et al., 1994); c) assess effects of study non-participation using information from WIC's database; and d) implement multiple imputation methods for missing data (Rubin, 1987, Schafer and Olsen, 1998). *If needed*, propensity score methods will be employed to re-match study groups, Propensity scores incorporate large numbers of covariates to yield quintile-based matching of experimental and control groups on scores. The scores are defined as the predicted probability of study group assignment obtained from logistic regressions that use baseline demographics and key outcome variables (Rosenbaum and Rubin, 1985). Such matching reduces "noise" in group comparisons, of particular value given the development nature of our proposed trial and small sample sizes. The propensity score method does not focus on effect estimation, limited here by the relatively small sample sizes and will be used only to closely examine any issues with randomization and attrition to help plan for a Phase II trial of e-SBI. *Given random assignment and the relatively homogeneous nature of the clinic population with respect to age, income and education, we do not anticipate having unmatched study groups.*

We will evaluate e-SBI effects by examining group differences in changes in outcomes between baseline and the 3-month (T2) and, separately, the 6-month follow-up (T3). Mean differences will be estimated separately for the experimental and control group, combined using weights based on sample size, and compared across study condition. Categorical variables will be analyzed using the longitudinal Generalized Estimating Equations (GEE) approach (Liang and Zeger, 1986, Zeger and Liang, 1986) with indicator variables included for each follow-up and an exchangeable error correlation structure..

Post hoc analyses will focus on changes in drink size and drink size corrected quantities of consumption and *between and within group differences in drinking at* T2 and T3.

Power considerations. We consider the worst case scenario with *n=65* per group expected at 3-month follow up (T3). Tests of equality of means or rates of outcomes across matched groups would have a power of .82 to detect medium effect size differences in 2-sided tests with α=.05 (Cohen, 1988), corresponding to a group difference in proportions of at least 15% (e.g., any 5 plus - 5% in the experimental vs. 20% in the control group). Power for tests for 6-month outcomes (*n=60* per group), would be .78.