Study Name:	WOMEN SHARE
NCT Number:	NCT03323086
Grant Title:	REDUCING ALCOHOL-RELATED HIV/STI RISK FOR WOMEN IN REPRODUCTIVE HEALTH CLINICS
Grant #:	R34AA023158
Date of Document:	March 28, 2018
Principal Investigator:	Michael P. Carey, PhD The Miriam Hospital Providence, RI 02906
Content:	Statistical Analysis Plan

Data Analyses Plan

For the feasibility data, we produced a Consort flow figure and calculated summary statistics to report the findings. For the acceptability data, we calculated summary statistics and identified illustrative quotes from the semi-structured interviews. Prior to analyzing the efficacy data, we used t-test and chi-square analyses to determine if randomization produced equivalent groups with respect to demographic characteristics and outcome measures.

For the efficacy data, we used inferential statistics to compare the two conditions to one another. Specifically, we used generalized estimating equations (GEE; Hardin & Hilbe, 2013) to examine the impact of the intervention on primary outcomes (i.e., typical drinks per week, maximum drinks per day, number of heavy episodic drinking days, alcohol-related consequences, number of sexual partners, and alcohol use before sex) as well as to model change over time. We used an intent-to-treat approach such that data from all women were used. We hypothesized an intervention-by-time interaction such that participants in the BI, but not the control, condition would reduce alcohol use and sex risk behavior over time. All analyses examined the impact of time (baseline vs. follow-up), intervention condition (control vs. BI), and the time-by-condition interaction controlling for age. When interactions were not significant, models were run without the interaction terms. Count variables (i.e., drinks per week, maximum drinks per day, heavy episodic drinking days, number of partners, drinks before sex) were analyzed using a Poisson distribution unless otherwise indicated. Alcohol consequences were analyzed using linear models.

For condomless sex, a GEE model could not be fit due to significant over-dispersion bounded by a large number of zeroes. Instead, a zero-inflated negative binomial model was chosen based on the (a) lower AIC and BIC values in the negative binomial versus Poisson models, and zero-inflated versus non-inflated models; and the (b) likelihood ratio test comparing zero-inflated Poisson and negative binomial models (lr test: $\chi^2(1) = 60.76$, $p < .001$) (Long & Freese, 2006). Further, a zero-inflated model is considered particularly appropriate for sexual health research, such as condomless sex (He, Tang, Wang, & Crits-Cristoph, 2014), given that some mechanisms responsible for zero counts in the data differ from

the traditional regression model. For example, women who are not engaging in any sexual activity will produce zero counts in a condomless sex outcome, no matter how they vary on other predictors, and are therefore considered to be structurally different (in statistical modeling terms) than women who are at-risk of engaging in condomless sex but do not do so. Outliers (z scores ≥ 3.29; 3 cases at baseline and 2 at follow-up) were re-expressed to the next highest value (Tabachnick & Fidell, 2007).