

## Statistical Analysis Plan

### Sample Size and Power Analysis

Using a chi-square testing framework and a significance level of 0.05 and assuming two-sided hypothesis testing, both primary and secondary outcomes included a large enough sample size to detect differences between groups. Our a priori group sample sizes (usual care: n=191; LTWP: n=224) provided 80% power to detect between-group absolute differences in proportions ranging from 6% to 12% in the primary study outcomes (for example, screened: 65.2% [usual care] versus 77.7% [LTWP]; screened positive: 21.7% versus 34.2%, respectively; referred: 12.5% versus 23.1%, respectively; and attended treatment: 2.1% versus 8.3%, respectively). Our group sample sizes among participants completing a screen (usual care: n=63; LTWP: n=221) provided 80% power to detect between-group absolute differences in proportions ranging from 11% to 18% in the secondary study outcomes (for example, screened positive: 17.5% [usual care] versus 35.4% [LTWP]; referred: 3.2% versus 15.3%, respectively; and attended treatment: 1.6% versus 12.3%, respectively). The usual care rates used in these calculations were based on data published from earlier work on the LTWP program.

### Data Analysis

We will assess the adequacy of randomization by comparing LTWP versus usual care (UC) on demographic, obstetric, and psychiatric baseline characteristics, using t-tests or Wilcoxon rank sum tests for continuous variables and chi-square tests or Fisher's exact tests for categorical variables, respectively. Descriptive statistics will be calculated for each screening method (LTWP versus usual care), including counts and percentages for participants who were screened, screened positive, referred to treatment, and attended treatment. Unadjusted differences in outcomes between people assigned to LTWP versus UC will be calculated using chi-square tests because all study outcomes are dichotomous. Relative risk (RR) ratios will be estimated and reported along with their 95% confidence intervals. All hypothesis tests will be two-sided with a type 1 error set at 0.05, and all analyses will be performed using SAS statistical software, version 9.4 (Cary, NC).