BASELINE EQUIVALENCE TEMPLATE INSTRUCTIONS

The information collected will be used by you and your RETA to assess the baseline equivalence of the treatment and comparison samples. The HHS Teen Pregnancy Prevention evidence review standards require that quasi-experimental studies and RCTs with high attrition at the level of assignment establish baseline equivalence. These studies must demonstrate equivalence of the analytic sample on measures of age or grade, gender, and race and ethnicity. Studies of youth 14 years or older must also establish baseline equivalence on the behavioral outcome(s) of interest.

In total, this workbook has six tabs: two instruction tabs, and four tabs that you will complete to assess baseline equivalence of different analytic samples. This tab contains detailed instructions on how to complete the tabs by providing references to columns and lines in the tabs. Additionally, there is a second tab that provides illustrative instructions; this tab contains a copy of the baseline equivalence table that also appears on the four analytic sample tabs, but with comment boxes providing the instructions for a given cell or group of cells. Both sets of instructions can be applied to each of the four remaining tabs.

The remaining four tabs include the baseline equivalence table for you to complete, one for a different analytic sample (sample with baseline data, sample with post-program data, sample with first follow-up data, sample with second follow-up data). While you may not have completed data collection for your final analytic sample, examining data for these partial (or full) samples can be useful to identify issues such as a potential lack of equivalence due to high overall or differential attrition, variables with high rates of missing data, or lower rates of sexual experience than you were anticipating. Identifying these topics early can help you prepare to deal with these issues in your analysis plan.

Create up to four different analytic samples in your data file, one for each of the survey data collections. Define your analytic sample as anyone who responded to that follow-up and has data on at least one of your behavioral outcomes at baseline. Note that in defining your analytic sample, you should pool data from all cohorts that have completed the relevant follow-up survey at the time you are completing this workbook.

Choose the correct tab to work in, based on the analytic sample you are assessing. If you did not collect data for a follow-up period, leave that respective tab blank. (For example, if you did not collect a post-program survey from both groups, you can leave the "Sample w postprogram data" tab blank.) The sample size will vary across tabs as you will likely have different numbers of youth who completed each survey. Sample sizes may vary within each tab as well as you may have different item non-response across measures.

In each tab, enter data only in the highlighted yellow cells and enter baseline data on the analytic sample for that time point. For example, if you collected *postprogram* data from 100 youth in the treatment group and 100 youth in the comparison group, and at *baseline* 50% of each group reported being sexually active, and at *postprogram* 65% of each group reported being sexually active, in the postprogram tab, you would report the rate of sexual activity as 50% for each group of 100 youth.

Below are two sections:
(1) Detailed instructions for completing the tabs. As the tabs have the same structure, these directions apply to all tabs. Like the organization within tabs, the instructions are split between Demographics and Behavioral Measures. Note: many of these instructions are repeated within the tabs as comments.
(2) Instructions on how to interpret the p-value calculations in the file.

Section	Directions
	Detailed instructions on entering data
	Demographics
Age or Grade (Rows 4-10)	Choose whether you want to analyze age using a continuous measure (years) or a categorical one (grade or age categories). YOU NEED REPORT ON ONLY AGE OR GRADE, NOT BOTH. This choice should match how you plan to include age in your final analytic models. Any study enrolling youth in only one grade should assess differences in age across the groups. **If you choose to analyze age in years,** in row 4 enter the (1) unadjusted mean, (2) standard deviation, and (3) sample size for the treatment and comparison groups in columns C-E and columns G-I, respectively. The file will calculate a t-statistic for the variable, and present the p-value in column U. **If you choose to analyze grade, or age using discrete categories,** follow these steps to complete rows 6 through 10: (1) Construct the categories as you plan to analyze them in your final analytic models. You may want to separate out each age/grade (for instance, 9th grade, 10th grade, and 11th grade) or combine some ages/grades together (for instance, Ages 10-11 and Ages 12-13), for instance if some ages/grades have small sample sizes. You do not need to use all the rows. (2) Enter those categories as the labels for Categories 1-5 in Column B. Enter the number of individuals in each category in the sample size columns (Columns E and I for the treatment and comparison groups, respectively). The file will automatically calculate the percentage each age/grade category contributes to the total sample in Columns C and G, which should sum to 100 percent. The file will calculate a chi-squared statistic for the grade variable and report the associated p-value in Column U.

Sex (Row 11)	Please construct a binary measure where female = one and otherwise the variable is zero. For the row labeled "female," enter two pieces of information: (1) The percentage of the sample that is female in Columns C and G for the treatment and comparison groups, respectively. (2) The total sample with data on gender for each group in the sample size columns (Columns E and I). Note: this is the total number of individuals with data on gender, **not the number of girls in the sample.** For example, if you have 100 people in your Treatment Group, 50% of whom are female, in Column C you should enter 50 and in Column E you should enter 100. The file will calculate a t-statistic and p-value for the sex variable.
Race (Rows 12-16)	The four categories in Column B are the recommended categories for your analysis, using a variable constructed from both race and ethnicity data. (NOTE: If the suggested categories included in the template do not fit the makeup of your sample, you can construct your own categories that better suit your study. Please relabel the categories in Column B as needed.) For each row, enter the number of youth in your sample in each category in Columns E and I for the treatment and comparison groups, respectively. The file will automatically calculate the percentages each race category contributes to the total sample in Columns C and G, which should sum to 100 percent. The file will calculate a chi-squared statistic for the race variable and report the associated p-value in Column U.
	Behavioral measures
Binary measures (Rows 18-22)	Use rows 18-22 for binary outcome measures (for instance, ever had sex or had sex in last 3 months). Before entering data in these rows, make sure you have cleaned up your baseline data on the measures: (1) Please convert all yes/no responses to yes = one and no = zero in your data file. (2) For some measures, you may only have data from a subset of youth who responded affirmatively to a previous question about another related activity. When reporting on measures that youth only respond to if they have engaged in another activity, make sure to impute zero values for those youth who were not asked to respond to a subsequent question because they had not engaged in the other activity. For example, a question about pregnancy may be asked only of youth who report they have had vaginal sex. Youth who have not ever had vaginal sex cannot have been pregnant or caused a pregnancy, so you can logically infer that those youth would have also responded "no" for pregnancy and assign a zero value for them. This should be reflected in the data file before calculating means or percentages of the conditional-response measures. Once you have completed the above steps, complete the following steps for completing the tab for each binary measure: (1) Enter the measure name in Column B. (2) Report baseline percentages (out of 100) of the incidence of the outcome in Columns C and G for the treatment and comparison groups respectively. (For example, enter 50, not 0.50, to indicate 50 percent.) (3) Enter the corresponding sample sizes for each group in Columns E and I. The file will calculate a t-statistic and p-value for each binary measure entered.
Continuous measures (Rows 23-27)	Use rows 23-27 for a continuous outcome measure (for instance, the number of times having had sex); enter the measure name in column B. Enter the baseline mean, standard deviation and sample size in Columns C-E and Columns G-I for the treatment and comparison groups, respectively. The file will calculate a t-statistic and p-value for each continuous measure entered.
	Interpreting the p-value calculations
p-values (Columns U-V)	The spreadsheet will calculate p-values for each variable as an assessment of statistical significance between the groups. (If the p-value is less than 0.05, then the differences between the two groups at baseline are statistically significant. If the p-value is greater than or equal to 0.05, then the differences are not statistically significant.) However, you should take into consideration other factors when assessing whether the groups are substantively different from one another. For example, the p-value calculation assumes individual-level assignment to condition and may not take into account aspects of your study design, such as clustering. Evaluators of RCTs or QEDs where assignment is at a level other than the individual can adjust the p-values for clustering using a statistical package should they desire. These adjusted p-values should be entered into Column V. (For the ultimate demonstration of baseline equivalence of your analytic sample in your final report, we will ask you to be make that adjustment.) You should also consider the magnitude of the differences between the groups, in addition to the statistical significance. Large differences between the two groups may not be statistically significant (for example, if the sample size is small) but are still important when considering whether your comparison group is an adequate representation of the treatment group in absence of the intervention. Assess baseline differences for all analytic samples of interest--the samples with data at each time period in which you collected data. Baseline equivalence may vary by sample as different youth may have completed each follow-up.