ATTACHMENT F

ANALYSIS PLAN FOR PAF IMPACT STUDY

ANALYSIS PLAN FOR PAF IMPACT STUDY

The PAF impact study plans to experimentally evaluate programs for expectant and parenting teens in two selected sites (see Attachment A). In these two sites, youth will be randomly assigned (either individually or in clusters) to a treatment group that receives the program being tested or to a control group that does not and data for the impact studies will be collected through surveys of youth. When feasible, random assignment will take place within blocks (strata) to reduce the probability of chance differences between the treatment and control groups with respect to important factors (for example, geographic location). In the two random assignment sites (California and Texas), program impacts will be analyzed with survey data collected at baseline and at 12 and 24 months after baseline. Impacts will be analyzed separately for each site.

Our analysis plan for the impact study has three main components: (1) an early analysis of baseline data, (2) a primary impact analysis of key behavioral outcome measures, and (3) exploratory analyses of secondary research questions. These are described below.

Baseline analysis in random assignment sites. As soon as baseline data collection has been completed in each site, we will begin preliminary analyses of the baseline data. We will use these analyses to describe the study sample in each site and compare it with the target population. We will also assess whether random assignment successfully generated treatment and control groups balanced on important baseline characteristics. To support this analysis, our baseline survey will collect key measures of demographics (such as age, gender, race, and ethnicity) and other personal characteristics (such as prior sexual experience) needed to describe the study sample and examine the equivalence of the treatment and control groups.

Primary impact analysis in random assignment sites. Impact analysis will begin after the completion of follow-up data collection in each site. With a random assignment design, unbiased impact estimates can be obtained from the difference in unadjusted mean outcomes at follow up between the treatment and control groups. However, we can improve the precision of the estimates by using regression models to control for covariates, especially baseline measures of outcomes. Regression adjustment can also account for any blocking variables used in conducting random assignment, or for any differences between the treatment and control groups in baseline characteristics that arise by chance or from survey nonresponse.

The empirical specification for the model will depend on the unit of random assignment. With random assignment of youth, our model can be expressed as

$$(1) \quad y_i = \beta' x_i + \lambda T_i + \varepsilon_i$$

where y_i is the outcome of interest for youth i; x_i is a vector of baseline characteristics; T_i is an indicator equal to one for youth in the treatment group and zero for youth in the control group; and ε_i is a random error term. The vector of baseline characteristics x_i will include demographic characteristics such as age, race/ethnicity, and baseline measures of variables that are highly correlated with outcomes. These baseline characteristics will be gathered on baseline surveys. The parameter estimate for λ is the estimated impact of the program.

If clusters, rather than individual youth, are the unit of assignment, the estimation must account for the correlation of outcomes among youth in the same cluster, as they will all be

randomly assigned as a single unit, and each sample member cannot be considered statistically independent. To account for this dependence, we can modify the previous regression model as

(2) $y_{is} = \beta' x_{is} + \lambda T_{is} + \eta_s + \varepsilon_{is}$.

The general structure of the model is the same, but now y_{is} is the outcome measure for individual i in cluster s (and similarly for the treatment status indicator, T_{is}, vector of baseline characteristics, x_{is} and the error term ε_{is}). Most important, the error term in Equation (2) accounts for the clustering of youth within clusters because of the inclusion of the cluster-level error term η_s—a cluster "random effect." If this error term is excluded, the precision of the impact estimates could be seriously overstated. As in Equation (1), the estimated impact of the program is λ. Equation (1) or (2) will be estimated separately for each primary outcome in each site. Weights will be created for each site to account for any differences in random assignment or sampling probabilities among study participants.

To control for multiple hypothesis testing (the increased chance of falsely identifying an impact as statistically significant when examining effects on many outcomes), we will limit the primary analyses for each site to a small set of key outcomes. In selecting these outcomes, we will rely on the program logic model and data needs table developed for each site. We anticipate that most of these outcomes will be measures of sexual risk behavior and its health consequences (pregnancy, STIs, or birth) and also educational attainment, though the exact outcomes selected will vary by site. Within this small set of key outcomes, we will also consider applying a formal statistical correction for multiple hypothesis testing.

To support these analyses, the follow-up surveys will include measures of all key outcomes —primarily pregnancies, births, sexual risk behaviors, and educational attainment. We will also include these measures and related measures on the baseline survey, so that we can include them as covariates in the regression models used to estimate program impacts.

Analysis of secondary research questions. In addition to our primary impact analysis, we will also define and answer additional secondary research questions for each site:

- ***Subgroup analyses***. To examine whether the programs were more effective for some youth than for others, we will estimate impacts for subgroups of youth by adding a term to Equations (1) and (2) that interacts the treatment indicator by a binary indicator of a particular subgroup. The regression coefficient on this term provides an estimate of the difference in the program effect across the subgroups. Subgroups of particular interest include race/ethnicity, and whether female was pregnant or parenting at baseline. To support these analyses, we will include these subgroup variables on the baseline survey.
- ***Impacts on mediating variables.*** In addition to primary analysis of program impacts on outcomes of most central importance, as part of secondary analysis we will also examine program impacts on key mediating variables specified in the program logic model for each site (for example, knowledge of contraception and attitudes about subsequent pregnancies). We will estimate impacts on these outcomes following the same approach described in Equations (1) and (2). These mediating variables will be drawn primarily from the short-term follow-up survey, which will be conducted 12

months after baseline. We will also include selected mediating variables on the baseline survey, to include as covariates in the regression models.

- ***Variation in impacts by participation levels.*** Our primary impact analysis will include the full study sample, yielding intent-to-treat (ITT) estimates that do not account for varying participation rates among youth assigned to the treatment group. As exploratory analyses, we will consider adjusting for participation levels in two ways. First, to account for youth who do not attend *any* program sessions or activities, we can make the standard Bloom adjustment to calculate estimates of the treatment on the treated (TOT). Second, to explore the association between program dosage—the degree of program participation—and impacts, we can conduct propensity score analyses, whereby youth with the highest program attendance are matched to a subset of control group youth with similar demographic and baseline characteristics. To support these analyses, our baseline survey will include a broad range of demographic and other personal characteristics to consider as potential matching variables.