Impact Study Analysis Plan

We will conduct an Intent-to-Treat (ITT) analysis of the outcomes. This ITT analysis provides an estimate of the impact of the LifeSet on those who were eligible for and offered the program. The ITT estimate is defined as the difference between the average outcomes for those randomized to the LifeSet program, the treatment group, and those randomized to the control group, adjusting for pre-randomization covariates. All eligible individuals randomized to the treatment group will be counted in the treatment group, regardless of whether they actually engage with LifeSet. All eligible individuals randomized to the control group will be counted in the control group, even if they enroll in LifeSet. Because the processes that lead a treatment group youth to drop out of the program, or a comparison group youth to seek and enroll in LifeSet, are non-random and likely related to study outcomes, assigning treatment and control conditions based on the original assignment in this way, provides the most unbiased estimate of treatment effect. The ITT estimate answers the question—what is the effect of LifeSet on youth eligible for the program—and because it reflects the drop-out and substitution common in social programs, is considered to be a particularly policy-relevant metric.

Calculation: The ITT estimate is measured as the average individual outcomes for the treatment group less the average individual outcomes for the control group. We control for pre-randomization covariates using a regression framework. Specifically, the ITT estimate, β^T, would be measured using the regression equation below:

$$Y_i = \alpha + \beta^T T_i + \sum_{n=1}^{N} \beta^n X_i^n + \varepsilon_i$$

Y_i is the outcome for each individual, *i*, that was randomly assigned. T_i is an indicator equal to 1 for individuals who were assigned to the treatment group and 0 for individuals assigned to the control group. β^T is the parameter of the ITT effect on the outcome (Y_i) the number of population members assigned to the treatment group and control group, respectively. X^n is a vector of pre-randomization covariates and β^n is the vector of coefficients on the covariate, X^n. ε is the regression error term. For continuous outcomes, we will utilize an OLS regression model and for binary outcomes, we will utilize a logit or probit model for estimation. The inclusion of the pre-randomization covariates is intended to improve the precision of the estimates.

The exact covariates will be finalized after reviewing the data for data quality and completeness. In addition, the sample will be evaluated for equivalence between the treatment and control groups on observable pre-randomization variables. Although random assignment is intended to create two equivalent groups, small samples can result in some differences between the groups by chance. Variables that show

differences between the two groups at p = .05, that is, with at least 95 percent confidence they are different, will be included as covariates in the regressions.

We anticipate conducting exploratory subgroup analyses of program impacts on substantively important subpopulations such as those defined by youths' gender, parenting status at baseline, juvenile justice history at baseline, and age at randomization. As we do not plan to stratify the randomization by any subgroup, all results of subgroup analyses should be considered exploratory in nature. Additionally, as sample sizes for subgroup analyses will be smaller than in the main analyses, our ability to detect effects of the program will be more limited.

Depending on the take up and crossover rates for the evaluation, we may also estimate the Treatment-on-the-Treated (TOT) estimate using an "instrumental variable" estimate (IV) (Angrist, Imbens, & Rubins, 1996). We would utilize this estimate if there was low take up among the treatment group, e.g. they enrolled but did not engage, or high crossovers among the control group, e.g. they were randomized into control but despite this received services from LifeSet. The IV estimate is a "per-person served" estimate, among those who comply with their random assignment that accounts for take up and the crossovers. For example, imagine that all study participants can be divided into three types of individuals: 1) those who will always engage in LifeSet regardless of whether they are enrolled in it or not; 2) those who will never engage in LifeSet even if they are enrolled in it; and 3) those who comply with whatever assignment they are given, whether it is to engage in LifeSet or to remain in the control group. The IV estimate represents the effect of LifeSet engagement on study outcomes among this third group, the compliers. In the special circumstance where decisions to comply or not are independent of the study outcomes, the IV estimate also represents the average treatment effect.

Missing Data

For the baseline youth survey, nonresponse bias will be analyzed using youth characteristics noted in the child welfare administrative data to determine if certain characteristics are significantly associated with nonresponse. These include youth demographics (age, race/ethnicity, sex, county of residence) and foster care experiences (age at removal, total number of removals, length of time in placement, out-of-home placement types, etc.). Item nonresponse is anticipated to be 5 percent or less. Item nonresponse will be analyzed to determine whether data are missing at random or if certain items are consistently missing using the child welfare administrative data as well as valid responses to survey items. If data are missing at random, the we will consider whether imputation techniques to account for any missing data will be useful and feasible based on how extensive the missingness is and the availability of variables useful for imputation. We may also use relevant administrative data sources to determine youths' statistics on pre-tests at baseline. If we decide to impute any missing variables, we will use Stata to run multiple

imputation.[1] If missingness exceeds 5 percent for any items, we will work with the subcontracted survey firm to improve question wording for the follow-up survey waves to prevent future missingness. Items that are missing 10 percent or more of the time will be dropped from the survey.

[1] Little, Roderick JA, and Donald B. Rubin. *Statistical analysis with missing data,* 2002.