

ATTACHMENT D
ANALYSIS PLAN YES INITIATIVE

ANALYSIS PLAN FOR IMPACT STUDY

The purpose of The Youth Engagement in Sports: Collaboration to Improve Adolescent Physical Activity and Nutrition (YES Initiative) is to rigorously assess the impacts of the program. The program includes an estimated 6,100 6th, 7th, and 8th grade students recruited across 14 YES Initiative organizations. The youth will be randomly selected into one of two study conditions within each site – (1) an intervention group participating in the YES Initiative programming; or (2) a control group that is not participating in the YES Initiative yet has access to community sports and recreation resources (i.e., a business-as-usual control condition). Thus, the study design is experimental. 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 related to secondary research questions. These components are described below.

Baseline analysis. As soon as baseline data collection has been completed at each site, we will begin preliminary analyses of the baseline data. The baseline analysis will describe the study sample and assess whether assignments 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 sex, race, and ethnicity) and other individual-level characteristics (such as level of physical activity) needed to describe the study sample and examine the equivalence of the treatment and control groups.

Primary analysis. An outcome evaluation will begin after the completion of follow-up data collection. Unbiased estimates of the program’s impact can be obtained with an intent to treat (ITT) difference-in-difference (DID) design in which the pre- and post-intervention change in outcomes among the intervention group are compared to the pre- and post-intervention change in outcomes among the control group. In addition to calculating the impact of the ask (ITT), an average treatment effect for the treated (ATT) or a complier average causal effect (CACE) design may be explored. An ATT design calculates the impact on those treated while the CACE calculates the impact on students who complied with their randomization assignment.

To estimate the difference-in-difference impacts of the YES Initiative, each outcome Y_i is modeled by the following equation

$$Y_i = \alpha + \beta T_i + \gamma t_i + \delta (T_i \cdot t_i) + \varepsilon_i$$

where the coefficients given by the greek letters α , β , γ , δ , are all unknown parameters and ε_i is a random, unobserved "error" term. The equation includes the following interpretation

α = constant term

β = treatment group specific effect (to account for average permanent differences between treatment and control)

T_i = 1 if in the YES Initiative, 0 if in the control group

γ = time trend common to control and treatment groups

t_i = time

δ = true effect of treatment

To support these analyses, follow-up surveys will include measures of key individual-level outcomes— sports participation, the number of sports activities, and amount of daily physical activity. Additional outcomes include physical literacy, as well as the percent of days consuming fruits/vegetables and sugar-sweetened beverages.

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

- **Subgroup analyses.** To examine whether the programs were more effective for some youth than for others, we may estimate outcomes for subgroups of youth by adding a term to the model 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 sex, race, ethnicity, and sports participation at baseline. To support these analyses, we will include these subgroup variables on the baseline survey.
- **Variation in impacts by program delivery.** Our primary impact analysis will include the full study sample that do not account for varying delivery of programming among youth assigned to the treatment group. As exploratory analyses, we will explore the association between program delivery—treatment setting, duration and intensity of treatment, type of intervention—and impacts. To accomplish this, we will conduct two separate analyses: (1) Estimate impacts separately by community (e.g. YES Initiative organization), and compare the magnitude of the impacts in communities by type of organization) (2) Estimate impacts by intensity of program. To support these analyses, we will combine the survey data used to estimate outcomes with the rich implementation data including type of organization, intervention program characteristics, and attendance as a means to quantify the delivery as a variable that might explain variation in outcomes/impacts.