

South Texas Early Prevention Studies PreK

Study Design and Protocol

September 28, 2018

Statistical Analyses

Summary statistics are presented as means (standard deviations; S.D.), and frequencies (percentages) to describe the treatment groups, and p-values adjusted by random factors. The primary aim of this report was to assess the longitudinal effect of the BN CSHP on the CRF of children as compared to a control group. The CRF was measured by the number of laps in the PACER fitness test. The counts of laps were positively skewed and assumed to follow a Poisson probability distribution. The linear fixed predictors of the average number of laps were time (continuous variable consisting of the periods of data collection), treatment effect (discrete factor with control group as the reference level), and their interaction (the nested district-level, school-level, and classroom-level to account for random intercepts).

A generalized linear mixed model (GLMM) with a natural logarithm link and a Poisson probability distribution was fitted to the data using the R package lme4, version 1.1-27.1.²⁷ Upon detecting a significant overdispersion, a GLMM with a negative binomial distribution was used instead to assess the unadjusted and adjusted effect of the intervention and to obtain 95% confidence intervals (CI).

The models were adjusted for covariates sex, age, body mass index (BMI), education level of caretaker who responded to the demographic survey, household income to poverty ratio (IPR), time watching T.V., time playing with electronic devices, hours of sleep at night, and time exercising outside school. The interactions between period and treatment by BMI, sex, and age were also included in the full model. As suggested by Brown²⁸ all continuous variables were transformed to a unit-scale and used the Bound Optimization BY Quadratic Approximation²⁹ optimizer in the lme4 package to ensure convergence. The analysis was performed in R version 4.1.0 [R Core Team, 2013].