GPS (Giving Parents Support): Parent Navigation After NICU Discharge

NCT02743472

Statistical Analysis Plan

Updated/approved by Children's National IRB: 2/15/17

PI: Karen Fratantoni, MD, MPH

Statistical Considerations:

We will conduct intent-to-treat analyses to assess effects of the PN intervention such that randomized families will be analyzed according to their assigned intervention group, regardless of departures from the randomized intervention. While it is possible that families may seek some type of parental support independently of the intervention, these potential effects may be evaluated in sensitivity analyses. Our analysis strategy will first compare baseline characteristics between the PN and control groups to evaluate the balance produced by randomization. It is expected that randomization with stratification on one key confounder of birth weight (<=1500 grams or >1500 grams) will result in a balanced distribution of factors between the PN and care resource groups on most, if not all factors. However, even with randomization some imbalance may occur on a few factors and it is important to identify such factors, as these are the ones that need to be controlled during analysis.

Aim 1a. and 1b. The outcome measures involved in this aim (e.g., parenting self-efficacy measured using the PMP S-E and perceived stress scale), are derived from multi-item scales and will be treated as continuous measurements. For each scale, the mean score after baseline will be assessed at each of the visits (1 week, 1 month, 3 months, 6 months and 12 months). The overall mean difference from baseline for each of the scales between the two intervention groups will be used to evaluate Aim 1. We will assess the overall change in scores for the self-efficacy and depression scales, i.e. the population averaged effect of the intervention on changes in self-efficacy and depression, using a generalized estimating equation (GEE) model to account for the correlation between the repeated change measurements among primary care givers. The GEE model is reasonably robust to missing data. We will use an exchangeable correlation structure, but in sensitivity analyses, will also test unstructured correlation and compare parameter estimates. The main independent variable in these analyses is intervention group, however we will also include (if necessary) a cross-product term of group-by-time which will enable us to estimate differences between the intervention groups over time. The group-by-time term will be included only if it is statistically significant. The model will include covariates reflecting the randomization stratification variables, and other covariates only if imbalances are observed in the randomization groups at baseline. In addition, we will use the model to plot the change in parental self-efficacy or parenting stress over time-by-group with 95% CI around each time point estimate. The degree of change at 3-months will allow us to assess the rapidity of any effect of intervention and a descriptive comparison of the differences between the groups at 6- and 12-months will allow us to assess the degree of persistence of any effect. We will consider the difference between groups to be statistically significant when the 95% confidence intervals around the difference do not include 0.

Aim 2a. and 2.b. will compare measures of anxiety using the STAI, and depression using the CES-D in primary caregivers from the intervention and control groups. The independent variable will be the intervention group in both models. If necessary, the models will include covariates known to impact parental depression and anxiety (if different between the groups) and stratification variables. We will estimate and compare the covariate adjusted mean difference in the scores between the intervention

groups using GEE models as described above (Aim 1). The difference will be considered to be statistically significant if the 95% CI around the difference fails to include 0.

As in Aim 1, group-by time-interaction models will be used to evaluate differences between the groups over time.

Aim 3. The NICU graduates may experience multiple ED visits or hospitalizations any time after baseline. The number of ED or hospitalization counts will be summed for each of the intervention groups. To assess differences between the intervention groups for each of the outcomes (ED visits or hospitalizations), we will use the Wilcoxon–Mann–Whitney two-sample rank-sum test, which is appropriate for use with nonparametric data, such as the count data from this aim. We will define the outcome, improved immunization status, as receipt of two diphtheria-tetanus-acellular pertussis (DTaP) vaccines, two Hemophilus influenzae b (HIB) vaccines, and two pneumococcal conjugate vaccines (PCV-13) by the end of the observation period (12 months after baseline). This binary outcome will be contrasted using an unpaired t-test with unequal variances to measure the hypothesis that the intervention results in improved immunization status.

Sensitivity Analyses. In sensitivity analyses, we will explore the effect of departures from assumptions made for the main analysis. A key assumption in our analyses is that missing data are missing at random, and groups who are and are not lost to follow-up are similar. We will compare characteristics of families who are, and are not, lost to follow-up. If significant differences between these two ($p<0.05$) are observed, we can test for intervention effects, but with the assumption that individuals with missing data have worse outcomes. We will then report these 'worst case' scenarios and report any changes in the clinical interpretation of the trial.

Sample Size & Power. We calculated power to detect projected differences between the intervention groups for the 3 primary outcomes. All computations are based on a projected sample size of 300 families of NICU infants (i.e. 300 primary care giver evaluations, and 300 NICU infant outcome evaluations) with randomization into 2 equal groups of 150, with data collected over multiple time points during the 12 months of follow-up post discharge. In Aims 1 and 2, we will have up to 5 repeated measures after baseline (1-week, 1-month, 3-months, 6-months, and 12-months), which increases our statistical power for those aims. We assumed a modest within-subject correlation of 0.50 between these repeated measures. Using Bonferroni correction, we adjusted alpha (0.05) for the number of primary outcomes, n=3. For Aim 1, we estimate that we will have excellent power (93%) to detect differences in parental self-efficacy scores as small as 0.33SD. This estimate is based on PMSPE score distributions as reported in the literature. For the main infant outcome of number of hospitalizations during the 12-month follow-up post discharge period, we estimate that we have excellent power (99%) to detect a difference of 1 in the median number of hospitalizations between the groups. Given our conservative assumptions, including modest correlation (r=0.5) between repeated measures and Bonferroni correction for the number of primary outcomes, we believe that our sample size ensures sufficient power to measure clinically relevant differences between the intervention groups.