THE HOUSTON HOME-BASED INTEGRATED INTERVENTION TARGETING BETTER ASTHMA CONTROL (HIITBAC) FOR AFRICAN AMERICANS

Statistical Analysis Plan

Revision Dated 9/9/2018

Power and Sample Size

Power and sample size calculations were performed using the PS program of Dupont and Plummer. The study was initially designed to recruit 300 patients in total. Based on the literature and our previous experience with the targeted participant population, we estimated a 35% LTF rate, resulting in 97 patients per group at the end of the study.

The primary outcome measures were the ACT™ score, MiniAQLQ score, and self-reported ED visits. For the ACT™ score, assuming a standard deviation of 4, a sample size of 97 participants per group gave greater than 99% power to detect a minimal clinically important difference (MCID)—which has been determined to be 3 for the ACT™ score—at the significance level of 0.05 (2-sided) using an independent, two-sample t-test assuming equal variances. For the MiniAQLQ score, assuming a standard deviation of 0.6, a sample size of 97 participants per group provided roughly 99% power to detect a MCID—which has been determined to be 0.5 for the MiniAQLQ score—at a significance level of 0.05 (2-sided) using an independent, two-sample t-test assuming equal variances. For ED visits, assuming a standard deviation of 3 for the number of self-reported ED visits for asthma over the previous 12 months, a sample size of 97 per group provided about 90% power to detect a difference of 1.4 in annual ED visits between the two groups at the significance level of 0.05 (2-sided) using an independent, two-sample t-test assuming equal variances.

Randomization

We used a permutated block design with varying block sizes, stratified by age (less than 55 and 55 years or older) and sex, to randomize enrollees into either the clinic-only or home-visit group. The randomization program was written into our online HIITBAC CTMS. Randomization was done at the end of the baseline clinic visit. No one knew the randomization status until the baseline visit was completed. Subsequently, the clinical and home-visit teams were aware of the randomization status, which was needed to provide customized standard of care. The persons

making the mid-year phone calls were blinded to randomization status. No data analysis aside from data audits, validity checks and summary data for the Data Safety Monitoring Board was begun until the intervention and data collection were completed.

Statistical Analysis

Summary statistics were calculated for the entire cohort, stratified by treatment group. Baseline demographics, clinical outcomes, and environmental exposures were summarized by means (standard deviations), medians (ranges), or frequencies (percentages). Baseline characteristics were compared between treatment groups using an independent two-sample t-test, Wilcoxon rank sum tests, chi-square tests, or Fisher's exact test. Baseline summary statistics were also compared between patients who completed the study (N=193) and those who did not (N=71). All hypothesis tests were two-sided and assessed at the 0.05 level of significance.

Primary Outcome Analyses. The primary outcome measures of interest for this study were the ACT™, MiniAQLQ, and ED visits in the previous 12 months at exit. The ACT™ score is the sum of the scores of the 5 items included in the questionnaire. If any of the five items was missing a response, then no score was computed. The MiniAQLQ scores were computed using two methods, per scoring recommendations. The MiniAQLQ was first computed as the average of the scores for the 15 items. If any of the 15 items was missing a response, then no score was calculated. Alternatively, the MiniAQLQ was also scored if 2 or fewer responses were missing by replacing the missing responses with the worst score (regression to the null). After replacement, the MiniAQLQ was then calculated as the average of the scores for the 15 items. The alternative MiniAQLQ was used as the primary measure for asthma quality of life. ED utilization was measured by the frequency of visits to an ED during the 12-month period prior to the exit visit.

The initial analysis of the outcome measures used an independent, two-sample, two-sided t-test, assuming equal variances, to compare the mean ACT™ scores, MiniAQLQ scores, and ED visits over the previous 12 months between treatment groups at study exit. The equal variance assumption was tested using the F test at the .05 level and, if rejected, Welch's t-test approximation was used as alternative analysis. Approximate normality was assessed by using histograms, box-plots, skewness, kurtosis, and the Shapiro-Wilks test. Departures from

normality were addressed by using nonparametric analytic methods or data transformation.

In addition to the t-test, the Wilcoxon rank-sum test and Fisher's exact test were used to compare treatment groups for the ED visits outcome measure. The Wilcoxon rank-sum test was used to test for significant differences in the location of the distribution of ED visits between treatment groups. The outcome measure was also redefined as a binary outcome to compare the proportion of participants who had at least one visit to the ED by treatment group using the Fisher's exact test. Other outcome measures were similarly compared between treatment groups using t-tests, Wilcoxon rank-sum tests, chi-square tests, or Fisher's exact tests.

To assess between-group differences for the primary outcome measures, we calculated the raw mean difference (MD) for the continuous variables, ACT™ and MiniAQLQ; and the odds ratio (OR) for the binary variable, ED visits.

Exploratory Analyses. Data on ACT™, MiniAQLQ and ED visits were collected at baseline, and approximately 6 months and 12 months after enrollment. As an initial exploratory analyses of our primary outcome measures, we performed a repeated measures analysis using a general linear mixed model (LMM) for the continuous variables (ACT™ and MiniAQLQ) and a logistic regression using a generalized estimating equation (GEE) for categorical data (ED visits) to examine change over time. These models assume an intent-to-treat analysis. The LMM and GEE models include all available observations in the analyses and assume the data are missing at random, applying the restricted maximum likelihood estimation method. For the repeated measures analyses, we included time, study group, and the interaction between time and study group. The slopes between baseline and 6 months and 6 months and exit were calculated for each treatment group. The difference between treatment groups was compared at each time point. For the ACT™ and MiniAQLQ scores, the LMM was used to compute the difference in means and 95% confidence intervals (CIs) for effect size between the two treatment groups between baseline and exit. The logistic regression model was used to estimate the odds of having at least one ED visit in the 12-month period before exit compared with the 12-month period before baseline. The model was used to compute the odds ratio (OR) and 95% CIs for the effect size between the two treatment groups between baseline and exit.

We also performed exploratory analyses of potential predictor variables. Independent

linear regression models were used to explore the association between various demographic, clinical and housing baseline characteristics and the primary continuous outcome measures, ACT™ and MiniAQLQ, adjusting for study group and baseline scores. Predictor variables were selected for inclusion in the final multivariable linear regressions based on statistical significance (generally alpha = .05), frequency of missing observations, clinical relevance, and model fit statistics such as the R-squared value. Residual plots and diagnostics were used to assess linear regression model assumptions. Independent logistic regression was similarly used to assess the association between baseline demographic, clinical and housing characteristics and exit ED visits (yes/no), adjusting for study group and baseline ED visits. A multivariable logistic regression model was similarly constructed for key predictor variables of ED visits at exit.

In addition, we performed a number of post hoc subgroup analyses including (1) enrollees who completed all five home visits; (2) enrollees who did not change addresses during the intervention; and (3) enrollees who exited before Hurricane Harvey. We also examined the heterogeneity of treatment effect (HTE) by looking at the interactions between age and treatment group, and gender and treatment group. Interaction terms were assessed in linear regression models. All HTE analyses were post hoc and exploratory in nature.

Analyses were done using SAS 9.4 (Cary, NC) and Stata 15 (College Station, TX).