Title: RVA Breathes: A Richmond City Collaboration to Reduce Pediatric Asthma Disparities

NCT #: NCT03297645

Document date: June 9, 2022

Document type: Statistical Analysis Plan

Statistical Analysis Plan
RVA Breathes

We will explore the data using descriptive statistics (e.g., frequency distributions, contingency tables) prior to any hypothesis testing and consider transformation as appropriate.

Aim 1: We hypothesize that children in the two active interventions will achieve reductions in healthcare utilization (e.g., asthma-specific ED visits, hospital admissions) and school absences, and an increase in controller medication use (calculated from refill data from insurers) at post-treatment and follow-up than children in the comparator condition. Our primary outcome, healthcare utilization, will be defined as a composite of the frequency of asthma-related inpatient and ED visits. A composite variable (i.e., sum score) that takes into account the frequency of asthma-related inpatient and frequency of ED visits will be calculated for each participant. At baseline, historical data will be obtained to ensure that we have captured utilization for the 9 months prior to the baseline visit (i.e., utilization before the intervention begins). Our primary endpoint is the change in mean rate of ED/hospitalization in 9 months before to 9 months after RVA Breathes across groups. Generalized Linear Mixed Models (GLMM) will be used to determine within person changes in healthcare utilization across time, and then differences in healthcare utilization between groups (e.g., intervention vs. control). Thus, our main outcome is change in healthcare utilization at the individual level as an average of individuals. We will also *explore* differences in outcomes between the two active intervention conditions [(1) asthma education + home remediation + school intervention vs. 2) asthma education + home remediation] to determine the value of adding a school-based component. Differences in healthcare utilization (and other outcomes) will again be tested using a Generalized Linear Mixed Model (GLMM). GLMM models will be fit to multiple types of dependent variables including, Gaussian/Normal, Binomial, Poisson and Negative Binomial. The GLMM model fit will include a between-subject factor (Group, representing the control group and the two intervention groups) and will also include a random effect for the cluster (school). These models will be modified in longitudinal assessments by the inclusion of a within-subject factor (Time [pre, post)) and the interaction between Group and Time variables. The inclusion of the interaction terms allows for a formal test of differing courses between the groups over time. To account for seasonality and the impact of COVID-19, we will include sine and cosine functions of $x = 2*\pi*(time\backslash max(time))$ covariates and COVID-19 impact as covariate in analyses. COVID-19 impact will be designated as March 13, 2020 and Covariate=0 for months before COVID-19 and Covariate=1 for months after COVID-19. We will adjust for multiple comparisons using the Holm-Bonferroni method. In addition to healthcare utilization, we will also consider changes in the secondary outcomes of asthma action plans and controller medication use between conditions.

Aim 2: We hypothesize that secondary outcomes (asthma control, symptoms, quality of life) of children in the intervention groups will be superior to those of the control group. These outcomes will be tested using the GLMMs described under Aim 1 utilizing the proper distribution and link functions for the outcome variable examined.

Aim 3: We hypothesize that intervention effects will be mitigated in families with caregivers with lower levels of asthma self-management skills at baseline, as well as in those with higher baseline levels of stress and depressive symptoms. These outcomes will also be tested using the GLMMs described under Aim 1 utilizing the proper distribution and link functions for the outcome variable examined that have been modified by the inclusion of an additional effects measuring caregivers' self-management skills, as well as stress and depressive symptoms. The inclusion of these covariates will allow us to adjust the intervention effects and increase the

generalizability of the intervention effects based upon caregivers' characteristics.

Sample size. From RVA Breathes baseline data, Group 1 children (full intervention: CHW + HH assessors + school) had a 9-month ED visit rate of 2.22 visits. Group 2 (CHW + HH, no school) had a 9-month ED rate of 1.89 visits, and Group 3 (control) had a 9-month ED rate of 1.68. We used recent published rates from a comparable study that considered changes in child asthma-related ED visits from 6 months prior to intervention to 6 months post-intervention (Johnston et al., 2019). In this study, the intervention group went from 2.0/6mo (pre) to 0.55/6mo (post) and the control group went from 1.3/6mo (pre) to 0.85/6mo (post). Thus, the intervention group had a 72% reduction ((2.0-0.55)/2.0 = 0.72) and the "control" group had a 35% reduction ((1.3-.85)/1.3 = 0.35).

Basing our numbers on these same rates, we would expect the following power estimates with 9-month follow up data: 14% for Groups 1 vs. 2, 97% for Groups 1 vs. 3, and 88% for Groups 2 vs. 3. Using alpha = .0167, we would assume average clusters of 10 in 24 schools for 240 enrolled participants. Allowing for 20% attrition, this would be 300 participants. Moreover, using a Poisson model for analyses, we would use a CVB (between-cluster coefficient of variation) instead of ICC. These analyses would use a CVB for G1 vs. G2 = 0.1331, for G1 vs. G3 = 0.1820, and for G2 vs. G3 = 0.0502.

Missingness. We will handle missing data in several ways. Health utilization is being reported by caregivers (self-report) and we are also receiving health encounter data from insurers and hospital billing systems. For families that have been lost to follow-up but have not asked to withdraw from the study, we will still receive and use their objective health encounter data. This will limit missingness related to our main outcomes when using encounter data. For self-report data, we will conduct a sensitivity analysis on raw data and examine patterns of missingness. We then will use a pattern-mixture model with multiple imputation from other available timepoints. We can then compare those analyses to analyses using objective health encounter data.

Missing Data and Participant Attrition Strategies:

The primary analysis will follow the approved statistical analysis plan, without any changes for missing data. After fitting this model, missing data patterns will be examined and their patterns will be examined, in particular looking for a monotone missing data pattern indicative of attrition. If, as expected, we find that there is a monotone pattern of missing data (i.e., once an individual misses an observation, they will be missing all subsequent observations), this will provide evidence that the missing completely at random (MCAR) and missing at random (MAR) assumptions may be violated. To examine the impact of monotone missing data, multiple imputation using a pattern-mixture model with control-based imputation will be utilized. We will create 50 imputed datasets for the outcomes and analyze them using the primary analysis model. The results from these 50 analyses will then be combined to derive an overall result. This will allow us to perform a sensitivity analysis to see how much bias the missing data causes.

Johnston, K., Meier, M., & Federico, M. (2019). A Community Health Worker led home visit program integrated into asthma specialty care decreases health care utilization and shows a sustained impact on asthma control. *Am J Respir Crit Care Med, 199:* A4061.