Official Title: Hospital to Home Outcomes (H2O): A Study to Improve the Fluidity of Transitions Between Hospital and Home
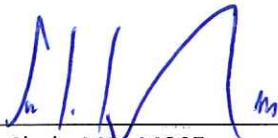
NCT: 02081846

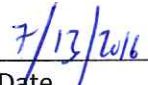**Date:** 20 September 2018

H2O Statistical Analysis Plan for AIM 3

Prepared by: Heidi Sucharew, PhD

Jane Khoury, PhD

Date: 7/13/2016

Analysis plan approved by:

_____   _7/13/2016_
Samir Shah, MD, MSCE   Date
Professor, Division Director
Division of Hospital Medicine
PI H2O

The results of H2O Aim 3 will be assessed as described in this document. This statistical analysis plan begins with an overview of the key components of the study including its design, objectives, and outcome measures. This is followed by a detailed description of the statistical analyses that will be utilized to address Aim 3 study hypotheses.

1. Overview of study: This study is single center, parallel, randomized trial to evaluate the effectiveness of a nurse home visit. Patients that meet eligibility criteria will be randomized 1:1 to receive either the intervention group (a nurse home visit within 96 hours of discharge) or the control group (standard-of-care). The randomization scheme utilized permuted block with two stratification factors, SES and service unit, to ensure equal allocation of the intervention within each stratification block. Random permutations within each block within each stratum were generated using a random number generator available in SAS Proc Plan. A total of 1500 patients will be enrolled in this trial.

1.1 Study Objective: To evaluate the effectiveness of the nurse home visit on health care utilization within 30 days of the index hospital discharge and family-defined transition outcomes defined in Aim1.

> Working hypothesis 1: Patients who receive the nurse visit will have a significantly lower occurrence of any unplanned re-hospitalization and/or any emergency/urgent care visits within 30 days of hospital discharge compared to the patients enrolled in the standard-of-care group.

> Working hypothesis 2: Patients who receive the nurse visit will have significantly better patient- and family-defined transition outcome measures compared to patients in the standard-of-care group.

1.2 Outcome measures

Primary

- Unplanned re-hospitalization and/or any emergency/urgent care visits within 30 days of hospital discharge.
    - Dichotomized as any event reported through the HIE – HealthBridge
    - Unplanned CCHMC hospital admissions will be defined as <=24 hour difference between scheduled entry and admission time for CCHMC events with this time data available
    - Non-CCHMC hospital admissions and outpatient visits will be classified as unplanned or planned by study team review of diagnosis data (ICD9/10) received from HealthBridge
    - All ED visits will be classified as unplanned
    - CCHMC urgent care visits will be classified as unplanned

Secondary

- Post-Discharge Coping Difficulty Scale (PDCDS) total score
- Child health questions, 6 individual Likert scale items scored 1 to 4 (NOTE: these questions were modified from the P-PAM and must be reviewed and approved by Insignia due to copyright restrictions prior to any publication/presentation of these questions)
- Knowledge of red flags (number correct)
- Days back to normal routine perceived by caregiver (number of days 0 to >=14 days)
- Type specific, hospital and ED admissions only, within 30 days of hospital discharge. (dichotomized as any hospital/ED reported event through the HIE or CCHMC EPIC)
- Early re-utilization within 14-day post-discharge (HIE data or CCHMC Epic or parent report)

NOTE on cost outcomes – Lenisa Chang will analyze cost data, both out of pocket and utilization cost, and is <u>not</u> included in this statistical analysis plan.

1.3 Subgroup analysis measures

- Financial strain subgroups as high strain vs. low/moderate strain
- Access to primary care subgroups as high vs. moderate vs. low access
- Caregiver education as high school graduate vs. no high school degree
- Publically insured (Paramount Advantage/Molina vs. all others)
- Service subgroups as hospital medicine vs. neurology/neurosurgery/yellow team

1.4 Schedule of visits

- Pre-discharge in person interview (baseline)
- Nurse home visit within 96 hours of discharge from index event for the intervention group only
- 14-day post-discharge phone call interview (14-day follow-up)

1.5 Data Management

CCHMC DMC staff will guide data management for this project. Please see data management plan in the appendix.

1.6 Interim Safety Analyses

No interim analyses are planned for this study.

2. Statistical Reports

2.1 Report Generation

2.1.1 Deviations from Statistical Plan

The final statistical reports will describe and justify any deviations from the original statistical plan described herein.

2.1.2 Software

The statistical reports will be prepared using SAS 9.3 and R 3.2 under Microsoft Windows operating system.

2.2 Analysis Populations

2.2.1 Screen Failures

Patients screened using EPIC data, randomly selected to be approached for participation, and did not give written consent are considered screen failures. Patients who give informed written consent but are not randomized are considered screen failures. Screen failures will be categorized as eligible but refused or eligible but not randomized (all eligible but not randomized for any reason except refused). The study coordinators will provide to the data management in excel format reasons for screen failure (categorized), age, census tract poverty (high vs. low), and service (HM vs. nuerology/neurosurgery/yellow team) for screen failed patients. These characteristics will be

summarized using counts and percentage for categorical variables and mean, median, variance, range for age for all screened failed patients. Note that this summary will be at the admission level, such that a patient can be included more than once for multiple admissions.

NOTE – the only data we have on those not randomized is the master file data which includes age, sex, language, unit/bed, service, address, diagnosis (free text a lot missing), active treatment team, pcp name, cvg 1 plan (payor would need to be classified).

2.2.2 Intent-to-Treat (ITT) Population

An ITT population will be defined as all patients who are randomized to a trial arm. For all practical purposes, it is expected that these patients will satisfy the minimum requirements for inclusion in the outcome analyses. The distinction between strict ITT patient and the ITT analysis group is made a priori to cover a situation in which a patient is randomized and, for whatever reason, the informed consent is invalid or withdrawn by the patient/family.

2.2.3 Intent-to-Treat Analysis (ITT-A) population

ITT patients with invalid/withdrawn consent will not be included in the statistical analyses, as we will not be able to collect/store any data from patients with invalid/withdrawn consent.

2.2.4 Per-Protocol Analysis (PPA) population

Patients in the ITT-A population who have no major protocol deviations/violations, as determined by the H2O study PIs, will be included in the PPA presentation and analyses. PPA will be defined by intervention adherence (received home visit within 96 hours for those in the home visit arm, did not receive a traditional home visit for those in the control arm within 96 hours), and completed (or partially completed) the 14 day follow-up survey within 14+7 days of index discharge date.

2.2.5 Completion

Patient will be considered to have completed the study once the patient completes the 14-day post-discharge phone call (+7 days).

2.2.6 Outliers

The data management and statistical analysis teams will identify patients having data values that appear to be potential outliers. These values will be sent to the research coordinators for verification. Clear identification of a value as an outlier will be based on medical judgment as well as on statistical grounds. In the event that outlier values are identified, any analysis using the actual values will be followed by an analysis that reduces the outlier effect.

2.2.7 Missing data

We will continue to collect EHR data (CCHMC Epic and HealthBridge) for the primary outcome for all study participants unless consent is withdrawn. Our approach for handling missing data will be to first assess consent withdrawal rates and evaluate whether or not the rates are very different between the two arms of the trial. If the rates/reasons are similar, we will perform the analysis assuming *Missing Completely at Random (MCAR)*, i.e., the propensity for a data point to be missing is completely random such that there is no relationship between whether a data point is missing and any values in the data set, missing or observed (missing data are a random subset of the data). If consent withdrawal rates are

very different between the two arms of a trial, then our results will be limited by concerns about potential bias.

For secondary outcomes of post-discharge coping difficulty, days back to normal, PPAM items, and red flags, subject-reported outcomes collected at the 14-day phone call, missing outcome data could occur for several reasons, such as item nonresponse, lost to follow-up, and data collection errors. Our approach for handling missing secondary outcome data will be to first assess missing patterns/reasons for missing, understand the distribution of missing data, and evaluate any differences between the two trial arms. This will allow us to assess whether an assumption of MCAR (missing data are a random subset of the data) is reasonable or if missingness is conditional on another variable in the dataset, i.e., Missing at Random (MAR). We will use descriptive statistics to compare baseline characteristics of patients with and without missing data. If MAR, we will incorporate variables that are identified to be related to the missingness in the analysis using multiple imputation, as multiple imputation works best when good predictors of outcome are available. Again, our first approach will be to monitor and minimize missing data, but if the amount of missing data seems likely to affect the study results, sensitivity analyses of the missing data assumption will be done using pattern-mixture models, as departures from the MAR assumption cannot be verified since missing values are not observed. We will evaluate if conclusions change under the pattern-mixture model structure.

2.2.8 Protocol Violations/Withdrawals

If a study participant is to be excluded/dropped from the study database and/or analysis, the PIs/coordinators will provide to the data management and statistical analysis team the following: 1) the specific reason for dropout (e.g., withdrawal, protocol violation), in as much detail as possible (such as failure to complete the intervention within the time window); 2) who decided that the participant would be excluded; and 3) whether the exclusion involves some or all types of participation.

2.3 Statistical Report Contents

The following section outlines the contents of the final statistical analysis. For all analyses, baseline is defined as the measurement prior and closest to the date of discharge from the index hospitalization event. Patients who are randomized will be analyzed in the groups to which they were allocated regardless of actual intervention received.

2.3.1 Description of study population

Sample size:
The total number of patients who are screened, consented, randomized, and withdrawn will be given. A summary of patients with protocol violations, such as failed nurse visit, or who were withdrawn will be provided.

Demographics and Clinical Characteristics, Patient and Parent:
Demographic and clinical characteristics as measured at baseline will be summarized by trial arm. Variables to be summarized are: child age, only child indicator, first hospitalization indicator, parent or caregiver age, race/ethnicity, education, marital status, household income, employment status, primary payor, financial strain, access to primary care, census track poverty (high vs. low percentage below poverty)service (Hospital Medicine vs. neurology/neurosurgery/yellow team). Primary discharge diagnosis ICD9 or ICD10 codes will be categorized using the AHRQ CCS algorithm. Variables will be summarized using descriptive statistics appropriate for each type of data item.

Baseline characteristics with large differences between groups will be included in analyses evaluating the effectiveness of the intervention.

2.3.2 Primary outcome analysis

Primary outcome: unplanned re-hospitalization and/or any emergency/urgent care visits within 30 days of hospital discharge

For the primary outcome, the dependent variable will be a dichotomized indicator of any occurrence of unplanned re-hospitalization and/or any emergency/urgent care visits within 30 days of hospital discharge recorded in CCHMC Epic or Healthbridge. Differences between intervention and control groups on this outcome will be evaluated using logistic regression with the stratification variables (census tract poverty and discharge service). Other potential covariates (e.g. baseline characteristics of age, first hospitalization, only child, caregiver education, with large differences between groups) will be included in a secondary analysis model evaluating the effectiveness of the intervention. In addition, a secondary analysis will be conducted excluding individuals enrolled in HNCC case management.

2.3.3 Secondary outcome analyses

- Post-Discharge Coping Difficulty Scale (PDCDS)

Post-Discharge Coping Difficulty Scale (PDCDS) will be measured during the 14-day post-discharge phone call. The PDCDS uses an 11 point scaling format (0-10). Higher scores represent greater coping difficulty. A total score will be computed as describe here:
    Items 8, 9, 10 a and 10 b will be reverse scored
    Total scale score = mean of 1+2+3+4+5+6a+7+8+9+10a+10b (missing items ignored) multiplied by 10 to give a score range of 0 to 110.
    Item 6b is included for comparison with 6a and will not be included in the total score.
The distribution of this outcome will be assessed for normality; if appropriate, we may apply a transformation or utilize a suitable link function. Differences between intervention and control groups on this outcome at 14-day post-discharge will be evaluated using a generalized linear model with the stratification variables (census tract poverty and discharge service) and other potential covariates included in the model.

- Parent-Patient Activation Measure (P-PAM), 6 individual Likert scale items scored 1 to 4 measured at baseline and 14-days post-discharge.
    - I know what each of my child's prescribed medications does
    - I am confident that I can tell when I need to go get medical care and when I can handle my child's health problem myself
    - I understand the nature and reasons for the illness that caused my child to be hospitalized
    - I know the different medical treatment options for my child's illness/condition
    - I know how to prevent further problems associated with my child's illness/condition
    - I am confident I can figure out solutions when changes occur with my child's health

Each item will be analyzed separately, i.e., a total score will not be computed. The distribution of each Likert-scale variable will be assessed for response range and skewness; if appropriate, we may dichotomize responses using the top-box approach, i.e., dichotomizing at strongly agree (4) vs less

than strongly agree (<4). Differences between intervention and control groups on this outcome at 14-day post-discharge will be evaluated using logistic regression (cumulative logistic for ordinal responses) with baseline PPAM item score and the stratification variables (census tract poverty and discharge service) and other potential covariates included in the model. *NOTE: these questions were modified from the P-PAM and must be reviewed and approved by Insignia due to copyright restrictions prior to any publication/presentation of these questions.*

- Knowledge of red flags defined as the number correct out of the number of flags given

The distribution of this outcome will be assessed; if appropriate, we may apply a transformation or utilize a suitable link function. Differences between intervention and control groups on this outcome at 14-day post-discharge will be evaluated using generalized linear model with the stratification variables (census tract poverty and discharge service) and other potential covariates included in the model. The number of flags given will be included as a covariate in the model.

- Back to normal routine perceived by caregiver - number of days 0 to >=14

The distribution of this outcome will be assessed for normality; if appropriate, we may apply a transformation or utilize a suitable link function. Differences between intervention and control groups on this outcome at 14-day post-discharge will be evaluated using a generalized linear model with the stratification variables (census tract poverty and discharge service) and other potential covariates included in the model.

- Secondary healthcare re-utilization measures:
  - Type specific, hospital and ED admissions only, within 30 days of hospital discharge. (dichotomized as any hospital/ED reported event through the HIE or CCHMC EPIC)
  - Early re-utilization within 14-day post-discharge (HIE data or CCHMC Epic or parent report)

These secondary healthcare re-utilization measures will be analyzed following the same approach described for the primary outcome.

2.3.4 Subgroup analyses

Exploratory prespecified subgroup analyses will be conducted by classification of subjects based on financial strain (high strain vs. low/moderate strain), access to primary care (high vs. moderate vs. low access), caregiver education (high school graduate vs. no high school degree), publically insured (Paramount Advantage/Molina vs. all others), and service (hospital medicine vs. neurology/neurosurgery/yellow team) then evaluate the effectiveness of the nurse visit within each subgroup using the approach described above for each outcome measure. Heterogeneity of treatment effects will be based on tests for interaction between subgroup level and treatment group and will be presented with effect estimates (including confidence intervals) within each level of the subgroup measure. The following outcomes will be evaluated: primary unplanned 30-day re-utilization and secondary outcome of days back to normal.

Financial/social strain- 10 categorical questions measured at baseline. A total financial strain score will be computed by summing the items with the top-box score for Likert items and yes for yes/no items. The distribution of the total scores will be obtained and dichotomized based on the top tertile.

- At the end of the month, do you generally end up with...? (Likert 1-3, 3 as top-box)
- During the past 12 months was there a time when you wanted to find work but were not able to?
- During the past 12 months, was there a time when your household did not pay the full amount of rent or mortgage?
- During the past 12 months, has your household not paid the full amount of any of the utility bills for electricity, heating or water?
- If you or your household had a problem with which you needed help (for example, child care, sickness or moving), how much help would you expect to get from family or friends living nearby? (in the same general region/city?) (Likert 1-5, 5 as top-box)
- Do you or your spouse/partner own your home? (reverse score - no)
- Do you or your spouse/partner own a car, truck or van? (reverse score - no)
- If you needed financial help during the next year, do you think you could count on people (family, friends, a bank or any source) to loan you $1000? (reverse score - no)
- During the past 12 months, have you borrowed money from friends or relatives because of financial circumstances or difficulties?
- During the past 12 months, have you or moved in with other people because of financial circumstances or difficulties?

Access to a primary care – measured at baseline using the Access subscale of the Parent's Perceptions of Primary Care (P3C) Scale. Responses are scored on a 5-point Likert-type scale (0=never to 4=always), with higher scores indicative of better primary care access for both routine and sick care. Items will be transformed to a 0 to 100 scale, with 100 being best, as follows: 0 = 0, 1 = 25, 2 = 50, 3 = 75, and 4 = 100. A total scale score will be computed as the mean of the nonmissing items. The total score will be categorized as always (100), almost always (75-99), and sometimes/often/never adequate (<75).

- Is it easy for you to travel to the doctor?
- Can you see the doctor as soon as you want for routine care (check ups, physicals) for the patient?
- If the patient is sick, can you see the doctor within one day?
- Can you get help or advice on evenings or weekends?

Caregiver education – measured at screening. Caregiver education will be dichotomized as high school graduate vs. no high school degree.

Publically insured – measured at screening. Publically insured will be dichotomized as Paramount Advantage/Molina vs. all others.

Unit - measured at screening using patient's discharge unit. Unit will be dichotomized as hospital medicine vs. neurology/neurosurgery/yellow team

Appendix: Data Management Plan

Data sources:
1) Interview data entered into REDCap (Research Electronic Data Capture)
2) EHR data (EPIC is the EHR system used at CCHMC) – labeled CSV files transmitted via SFTP
3) HealthBridge – labeled CSV files transmitted via SFTP
4) Heath Council (subsidiary of HealthBridge) Cost data
5) HNCC Care Management participation
6) Home Care Epic database
7) Readmission risk stratification data (data coming from Kate Rich)

- Interview data at discharge and 14-days, including patient medical record number, will be entered into REDCap and stored on SQL secured servers hosted by CCHMC's Biomedical Informatics (BMI) division, with state of the art security features including audit trails and password protection.

- For patients enrolled in Aim 3, select CCHMC EPIC data fields, including patient demographics, index admission characteristics, readmission, revisit data within 30-days of discharge, and nurse home visit characteristics (for the intervention arm only), will be extracted from the EHR EPIC system by CCHMC BMI staff and sent to CCHMC Data Management Center (DMC).

- HealthBridge routinely receives inpatient and emergency department admission and discharge information from hospitals across the Greater Cincinnati-Dayton community. As of January 1, 2015, all the local hospital systems (except for VA) were part of HealthBridge, which includes the hospital based free standing emergency units/urgent care. Using a patient panel provided by CCHMC DMC of patients enrolled in Aim 3, HealthBridge will match and track admission information for these patients. Healthbridge has an automatic patient match rate of 98.2%. Of the remaining 1.8%, 99.5% of those have suggested matches that are staff reviewed and linked. The remaining records are reviewed and manually matched by staff. A complete file will sent to CCHMC DMC every 6 months.

- The data sources will be linked together using CCHMC patient medical record number. The statistical software SAS will be used to import and merge data sources and to create analysis datasets. This will be done by CCHMC DMC.

- We will use guidelines set by Clinical Data Acquisition Standards Harmonization (CDASH) version 1.1, whenever possible for data collection fields. This standard guideline includes recommended data collection fields for 18 domains; including demographics, adverse events and other common domains that are common to most therapeutic areas and phases of clinical research.

- Ongoing data quality reports will be generated by the study data manager and will include data completeness, measured as the percentage of required variables completed divided by the expected required fields. In addition, descriptive statistics will be run to identify obvious errors in the database.

- During data collection, all associated research data will be physically stored on a password-protected secure sever maintained by CCHMC, except HealthBridge, which will follow their established standards until data is transferred to CCHMC. No data will reside on portable or laptop devices.

- Data collected during this study will be archived by CCHMC DBE in a study specific virtual archive folder.