

STATISTICAL ANALYSIS PLAN

VIDEO-BASED INTERVENTION TO REDUCE TREATMENT AND OUTCOME DISPARITIES IN ADULTS LIVING WITH STROKE OR TRANSIENT ISCHEMIC ATTACK (VIRTUAL)



Jose-Miguel Yamal, PhD
Charles Green, PhD

*Coordinating Center for Clinical Trials
Department of Biostatistics and Data Science
UTHealth School of Public Health
Houston, Texas, 77030*

TABLE OF CONTENTS

1	Statistical Analysis Plan and Statistical Reports.....	1
2	Study Synopsis.....	1
3	Outcomes	1
3.1	Primary Efficacy Outcome(s)	1
3.2	Secondary Efficacy Outcome(s).....	1
3.3	Additional Outcome(s)	2
3.4	Safety Outcomes	2
4	Definition of target population and study samples.....	2
4.1	Target Population	2
4.2	Intent-To-Treat Sample	2
5	General Statistical Considerations	3
5.1	Patient Accountability	3
5.2	Randomization.....	3
5.3	Blinding	3
5.4	Multiplicity	3
5.5	Monitoring and Evaluation of Clinical Center Effect	3
5.6	Baseline Descriptive Statistics	3
6	Primary Analysis	4
	Statistical Hypothesis	4
6.1	Primary Efficacy Analysis	4
	Missing Data	4
6.2	Interim Analysis for Safety.....	4
6.3	Secondary AnalysEs of Efficacy	5
6.3.1	Heterogeneity of the Primary Outcome Treatment Effect	5
6.3.2	Analysis Adjusting for potential confounders.....	6
6.3.3	Per-protocol Sensitivity Analysis	6
6.3.4	Analysis of Secondary Efficacy Outcomes after Trial Completion.....	6
6.4	Analysis of Exploratory Outcomes after Trial completion	10
7	Sample Size Justification	10
8	Safety Analyses.....	11
8.1	Overview	11

8.2	Definition of Adverse Events	11
8.2.1	Un-Blinded Statistical Monitoring	11
9	References.....	Error! Bookmark not defined.

1 STATISTICAL ANALYSIS PLAN AND STATISTICAL REPORTS

This document provides the details of statistical analyses planned for the VIRTUAL trial, including interim analyses for safety. In addition, it discusses the statistical issues relevant to these analyses (e.g., sample data to be used, missing data, adjustments for multiplicity, etc.).

The CCCT generates two statistical reports – an open report to be distributed to the VIRTUAL Trial Executive Committee and the Study Monitoring Committee (SMC), and a closed report to be distributed only to the SMC. The timing of these reports is determined in consultation with the SMC. Each report provides cumulative summary statistics on enrollment; subject status in the study (e.g., number completed each study visit assessment); baseline characteristics; protocol violations; data management/quality information (e.g., timeliness and completeness of data entry by the Clinical Sites; number of data queries generated and resolved); and safety data, including AEs and SAEs by AE code and severity. The statistics are provided for the overall study as well as by clinical center, where appropriate, in the open report. For the closed report only, the statistics are also provided by treatment group (A vs B).

2 STUDY SYNOPSIS

This is a two-arm randomized comparative effectiveness trial of a 6-month multidisciplinary telehealth intervention to test the impact of the intervention on key clinical outcomes, relative to standard care, among stroke survivors. The intervention is not blinded but the primary outcome, 24-hr ambulatory systolic blood pressure at 6 months after hospital discharge, will be analyzed blinded to the randomization group. This is considered a prospective randomized open, blinded end-point (PROBE) study.

3 OUTCOMES

3.1 Primary Efficacy Outcome(s)

- a. The primary efficacy endpoint is 24hr BP control according to ABPM 6 months post-hospital discharge, defined as < 125/75 mmHg.

3.2 Secondary Efficacy Outcome(s)

Secondary outcomes include:

- a. 24hr systolic blood pressure (BP) assessed with ambulatory BP monitoring (ABPM) 6 months after hospital discharge.
- b. 24hr Diastolic BP according to ABPM 6 months after hospital discharge.
- c. 24hr Systolic BP according to ABPM 12 months after hospital discharge
- d. 24hr Diastolic BP according to ABPM 12 months after hospital discharge.
- e. 24hr BP control according to ABPM 12 months post-hospital discharge, defined as < 125/75 mmHg.
- f. Composite recurrent vascular event rate (stroke, vascular death, non-elective coronary revascularization, myocardial infarction, heart failure hospitalization) 12 months after hospital discharge

- g. Proportion of uninsured stroke survivors who have obtained insurance coverage (including Medicaid, state programs, or commercial insurance) by 3- and 6-months following hospital discharge
- h. Hospital readmission rates and the number of acute healthcare visits (emergency department, urgent care) 3 months, 6 months, and 12 months after hospital discharge

3.3 Additional Outcome(s)

Other outcomes include:

- a. Daytime (6am-10pm), night time (10pm-6am) average systolic and diastolic blood pressure at 6 and 12 months.
- b. Proportion of current smokers who stop smoking or have at least one quit attempt by 6 and 12 months
- c. Proportion of smokers who have been treated for tobacco use disorder at 6 and 12 months
- d. Mean change in depressive symptoms using patients health questionnaire-9 (PHQ-9) at 6 and 12 months
- e. Adherence to the VIRTUAL intervention of telehealth visits (among those randomized to the VIRTUAL arm), defined as at least 2 of the telehealth visits attended
- f. Adherence to the VIRTUAL intervention of BP measurements (among those randomized to the VIRTUAL arm), defined as the number of weeks that have met the minimum number of 3 BP measurements per week.

3.4 Safety Outcomes

Safety assessment case report forms (CRFs) will be used to evaluate factors associated with aggressive BP reduction including syncope, falls, and new impaired renal function. These safety endpoints will be monitored by the SMC throughout the trial.

4 DEFINITION OF TARGET POPULATION AND STUDY SAMPLES

4.1 TARGET POPULATION

The target population for the VIRTUAL Trial is adult stroke survivors with presence of hypertension (by clinical history or hospital BP $\geq 140/90$ mmHg on two occasions) that have been discharged home after their index stroke. Other eligibility criteria are described in the protocol.

4.2 INTENT-TO-TREAT SAMPLE

The primary efficacy analysis is conducted according to Intention-to-Treat (ITT). Therefore, the primary efficacy analysis sample includes all subjects randomized. All randomized subjects will be analyzed according to the randomly assigned treatment arm, regardless of adherence to the assigned treatment. Note that some patients will likely be consented during their hospitalization but then not randomized if they were not discharged home and these will not be included in the ITT sample.

5 GENERAL STATISTICAL CONSIDERATIONS

5.1 PATIENT ACCOUNTABILITY

A flowchart will be created to present a summary of participant status. This flowchart will list the number of patients who were randomized to a treatment group. Then, within each treatment group, it will list the numbers of patients who crossed over from one arm to another, and the number of patients who completed the study, withdrew consent, and became lost to follow up (LTFU).

5.2 RANDOMIZATION

Randomization will be conducted centrally within the REDCap system. Using a permuted block randomization, with random block sizes, we will randomize within each site (up to 3 additional sites will come onboard), insurance status (any versus none), and cognitive score (Brief Neurophysiological Screening <9 versus 9+).

5.3 BLINDING

The study is conducted in a single-blind manner. The primary outcome, the 6-month ABPM, will be collected using a blood pressure device. The data, along with metadata regarding the times that the individual was asleep, will be analyzed by a physician blinded to the randomization group and otherwise not involved in this study. They will make the determinations of whether some BP measurements should be removed as treated as noise. The adjudicated data will then be used to the mean summary measure for the primary outcome.

5.4 MULTIPLICITY

For secondary outcomes, we will not account for multiplicity as these results will be treated only as supportive evidence (or lack thereof) of the treatment effect. The focus of reporting will be on point estimates of the treatment effect and corresponding confidence intervals.

5.5 MONITORING AND EVALUATION OF CLINICAL CENTER EFFECT

As other sites are onboarded, we will examine baseline and outcomes for each site and for all sites combined. Using tables and visualization, we will determine whether any sites are uniquely different from the overall group, considering the variability due to the number of subjects enrolled at that site (e.g. Funnel Plots).

5.6 BASELINE DESCRIPTIVE STATISTICS

A description of the baseline characteristics of trial participants will be presented by treatment group. Dichotomous variables will be summarized as number (%). Percentages will be calculated based on the number of participants with available data for that variable. Continuous variables will be summarized by the mean and standard deviation (SD) or median and interquartile range. In the case of variables with missing values, the denominator will be stated in the summary table or in a footnote to the summary table.

6 PRIMARY ANALYSIS

STATISTICAL HYPOTHESIS

6.1 PRIMARY EFFICACY ANALYSIS

Hypothesis 1.a: At 6 months, rates of controlled BP, according to ABPM, will be higher in the intervention arm compared to the standard care arm (“condition” in the model). Generalized linear modeling will evaluate BP control at 6 months as a function of treatment after adjustment for stratification variables (Section 5.2; site, insurance status and neuropsychological score). For each individual i :

$$\text{Logit (Probability of BP Control)}_i = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{Condition}_i$$

MISSING DATA

All subjects will be followed for 6 months for the primary efficacy outcome. Extensive efforts will be put forth to ensure complete follow-up, in particular with the ascertainment of primary efficacy outcome events. Nevertheless, missing data are anticipated with any longitudinal post-stroke study. If the ABPM is not acquired, we will impute missing values using any blood pressure measurements obtained within the collection window (e.g. with prioritization for measurements with higher correlation with the ABPM measurement). For example, if the ABPM is missing but the TrueBP measurement or other clinic visit measurements are obtained, the ABPM will be imputed using the TrueBP value and another field will capture whether the value is imputed or based on the ABPM). In the absence of these data points, and to the degree that the assumption of missing at random is plausible, analyses will address further missingness via multiple imputation.¹ Non-ignorable missing data patterns will be addressed through pattern-mixture modeling methods.² Sensitivity analyses will evaluate robustness of analytic conclusions to missing data.

6.2 INTERIM ANALYSIS FOR SAFETY

No interim analyses will be conducted for efficacy and futility. One formal safety interim analysis of recurrent vascular events (composite, adjudicated, but regardless of whether the events were adjudicated as related or not) will be conducted after enrollment of 267 total participants, of which about half will be randomized to the intervention group. Because ascertainment of safety outcomes may differ by group due to the increased touch points with the intervention participants (with possibly undercounting of events in the control group), we will compare the intervention arm only to assess whether recurrent vascular events ($p_{intervention}^{RV}$) occurred at a high rate, defined as above 20% as expected in the literature.³ The Bayesian model will assess the posterior probability $P(p_{intervention}^{RV} > 0.20)$, logistic regression will estimate the probability of recurrent vascular events in the active treatment after adjusting for stratification variables assuming a priors in the log-form $\sim \text{Normal}(\text{Mean} = 0, \text{SD} = 100)$ for the intercept and $\sim \text{Normal}(\text{Mean} = 0, \text{SD} = 10)$ for the treatment coefficient. The study may be discontinued if that probability is above 90%.

Probability of Recurrent Vascular Event		Estimated Probability of Identifying a Probability of Recurrent Vascular Events in Treatment > 0.20 as a Function of Posterior Probability Threshold (K = 1000 Monte Carlo Simulations)			
Treatment	Comparison	≥ 95%	≥ 90%	≥ 85%	≥ 80%
0.20	0.20	0.04	0.08	0.12	0.18
0.21	0.20	0.08	0.16	0.23	0.28
0.22	0.20	0.13	0.21	0.31	0.38
0.23	0.20	0.19	0.31	0.40	0.48
0.24	0.20	0.30	0.42	0.51	0.59
0.25	0.20	0.38	0.52	0.62	0.69
0.26	0.20	0.48	0.62	0.70	0.77
0.27	0.20	0.61	0.74	0.81	0.85
0.28	0.20	0.71	0.81	0.87	0.90
0.29	0.20	0.78	0.88	0.92	0.94
0.30	0.20	0.83	0.91	0.94	0.96

6.3 SECONDARY ANALYSES OF EFFICACY

6.3.1 Heterogeneity of the Treatment Effect for Blood Pressure Outcomes

In accordance with the NIH guidelines on the inclusion of women and minorities as subjects in clinical research, the primary outcome will be analyzed for evidence of differential treatment effects in subgroups. One of the aims of VIRTUAL is to assess whether at 6 months the impact of the intervention will be greater in Black and Hispanic SS (**Hypothesis 1.b**). Other pre-specified subgroups will include gender and sex as a biological variable.

A Bayesian multiple generalized linear regression will be fit with outcome 6-month BP control with variables the VIRTUAL arm and stratification variables. In addition, separately for each subgroup analysis, we will include covariates for the subgroup and the interaction between the subgroup and the treatment. The interaction term will be used to test for heterogeneity of the treatment effect across subgroups, and the subgroup-specific treatment effects and corresponding credible intervals will be constructed to evaluate subgroup effects.^{4,5} Similar analyses will be conducted for secondary outcomes of 6-month diastolic and systolic BP. For systolic and diastolic BP outcomes, the baseline value of the outcome will also be included in the model. For example, for 6-month systolic BP, the interaction will be adjusted for the baseline value of systolic BP.

Because the trial is powered to detect a main effect of treatment, rather than a clinically relevant interaction effect related to these subgroups, a clinically relevant interaction will be interpreted regardless of statistical significance.

Analysis adjusting for site

As a secondary analysis of the primary outcome, a random effect for site will be included in the analysis model in order to account for potential heterogeneity between sites. Should more than one site contribute observations to the trial, analyses will address this using fixed and random effects as necessary based on convergence diagnostics and Akiake Information Criteria.

6.3.2 Analysis Adjusting for potential confounders

Preliminary data analyses will characterize the sample in terms of demographic and baseline variables to assess treatment group comparability. Evaluation of potential confounding effects will examine demographic and baseline covariates exhibiting a correlation with both the treatment condition and outcomes^{6,7} by modeling the respective outcome with and without the potential confound. If the inclusion of a potential confound does not alter the statistical conclusion, the results will present the most parsimonious model. If inclusion of the potential confound does alter statistical conclusions the results will present both to permit readers to evaluate the evidence. Results will report all variables evaluated for confounding.

6.3.3 Per-protocol Sensitivity Analysis

In addition to the intent-to-treat analysis (primary), we will also conduct a per-protocol analysis with the subset of participants demonstrating adherence to protocol according to visit attendance (2+/4) and BP monitoring (>50% of weeks with 3 or more BP measures).

6.3.4 Analysis of Secondary Efficacy Outcomes after Trial Completion

Hypothesis 1.b: Parallel analyses will evaluate secondary outcomes systolic and diastolic BP at 6 months in the same way as discussed for secondary analyses of the primary outcome (Section 6.1), except linear regression models for continuous outcomes will be used instead of a logistic regression. In addition, we will test the following hypotheses.

At 6 months, the systolic BP according to ABPM will be lower in in the intervention arm compared to the standard care arm. General linear modeling will evaluate systolic BP at 6 months as a function of treatment after adjustment for baseline systolic BP and stratification variables (Section 5.2; site, insurance status and neuropsychological score). For each individual i :

$$\text{Systolic BP at 6 Months}_i = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{Baseline Systolic BP}_i + \beta_5 \times \text{Condition}_i + \epsilon_i$$

Where $\epsilon_i \sim \text{Normal}(\mu = 0, \sigma^2)$ and *Condition* is an indicator of being randomized to the intervention group (versus control).

At 6 months, the diastolic BP according to ABPM will be lower in in the intervention arm compared to the standard care arm. General linear modeling will evaluate diastolic BP at 6 months as a function of treatment after adjustment for baseline systolic BP and stratification variables (Section 5.2; site, insurance status and neuropsychological score). For each individual i :

$$\text{Diastolic BP at 6 Months}_i = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{Baseline Systolic BP}_i + \beta_5 \times \text{Condition}_i + \epsilon_i$$

Where $\epsilon_i \sim \text{Normal}(\mu = 0, \sigma^2)$ and *Condition* is an indicator of being randomized to the intervention group (versus control).

If the proposed models fail to meet statistical assumptions subsequent analyses will evaluate alternative parametric distributions (e.g. T-Distribution, Skew Normal, Gamma, etc.) as well as bootstrapped estimates. Final model selection will rely on AIC and inspection of residuals.

Similar models will be conducted for BP control, and systolic and diastolic BP at 12 months (hypothesis 1.c) as for the 6 month outcomes.

Hypothesis 1.c: At 12 months, the systolic and diastolic BP will be lower in the intervention arm relative to standard care. Moreover, the proportion of subjects with controlled BP according to ABPM will be higher in the intervention arm compared to the standard care arm. Generalized linear models will evaluate each 12 month outcome (systolic BP, diastolic BP and BP control) adjusting for baseline systolic BP (for systolic BP outcome) or diastolic BP (for diastolic BP outcome as well as stratification variables. Regarding analyses of continuous systolic and diastolic BP, for each individual i at 12 months:

$$\text{Systolic BP at 12 Months}_i = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{Baseline Systolic BP}_i + \beta_5 \times \text{Condition}_i + \epsilon_i$$

Where $\epsilon_i \sim \text{Normal}(\mu = 0, \sigma^2)$.

and

$$\text{Diastolic BP at 12 Months}_i = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{Baseline Diastolic BP}_i + \beta_5 \times \text{Condition}_i + \epsilon_i$$

Where $\epsilon_i \sim \text{Normal}(\mu = 0, \sigma^2)$.

If the proposed model fails to meet statistical assumptions subsequent analyses will evaluate alternative parametric distributions (e.g. T-Distribution, Skew Normal, Gamma, etc.) as well as bootstrapped estimates. Final model selection will rely on AIC and inspection of residuals.

$$\begin{aligned} \text{Logit (BP Control at 12 Months}_i) \\ = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i \\ + \beta_3 \times \text{Condition}_i \end{aligned}$$

Hypothesis 1.d: At 12 months, recurrent vascular event rates will be lower in the intervention arm as compared to standard care. Logistic regression will evaluate the probability of SS recurrent vascular event rate at 12-month follow-up as a function of treatment adjusting for stratification variables. Regarding analyses of the proportion of recurrent vascular events, for each individual i :

$$\text{Logit (Probability of Recurrent Vascular Event at 12 Months}_i) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Condition}_i$$

Hypothesis 2.a: At 3 and 6 months, a higher proportion of uninsured SS will have obtained insurance as compared to standard care. Logistic regression will evaluate the probability of SS having insurance at 3 and 6 months as a function of treatment adjusting for stratification variables. Regarding analyses of the proportion of participants obtaining insurance, for each uninsured individual i :

$$\text{Logit (Probability of Being Insured 3 Months}_i) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_3 \times \text{Condition}_i$$

and

$$\text{Logit (Probability of Being Insured 6 Months}_i) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_3 \times \text{Condition}_i$$

Hypothesis 2.b.1: At 12 months, acute care visit rates will be lower in the intervention arm as compared to standard care. We will use generalized linear modeling to evaluate the rate of acute care visits in SS (as measured by number of ED visits + urgent care visits) by 12 months follow-up as a function of treatment after adjusting for systolic blood pressure at baseline and stratification variables and an offset for the log of exposure amount (to account for differential follow-up time for those that are lost to follow-up or died before 12 months). For each individual i :

$$\ln(\lambda_i \text{ at 12 Months}) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Condition}_i$$

where $\text{Number of Acute Care Visits at 12 Months}_i \sim \text{Poisson}(\lambda_i \text{ at 12 Months})$.

If the model fails to meet Poisson assumptions, analyses will use alternative strategies including application of scaling coefficients, robust estimators, and/or evaluation of the fit of negative binomial as well as zero-inflated and truncated distributions as indicated by the data.

Hypothesis 2.b.2: At 3, 6, and 12 months, the probability of readmission will be lower in the intervention arm as compared to standard care. We will use logistic regression analysis to evaluate the probability of hospital readmission in SS at 3, 6, and 12 months follow-up as a function of treatment adjusting for systolic blood pressure at baseline and stratification variables. For analysis of the probability of readmission at three months, for individual i :

$$\text{Logit (Probability of Readmission by 3 Months}_i) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Condition}_i$$

For analysis of the probability of readmission at six months, for each individual i :

$$\text{Logit (Probability of Readmission by 6 Months}_i) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Condition}_i$$

For analysis of the probability of readmission at twelve months, for each individual i :

$$\text{Logit (Probability of Readmission by 12 Months}_i) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Condition}_i$$

A Poisson model will also be fit with an offset term for the log of exposure amount to account for differential amount of follow-up. If the model fails to meet Poisson assumptions, analyses will use alternative strategies including application of scaling coefficients or robust estimators, evaluation of the fit of negative binomial as well as zero-inflated and truncated distributions as indicated by the data.

Hypothesis 2.c.1: By 12 months, the intervention will demonstrate differential effects on acute care visit rates (as measured by number of ED visits + urgent care visits) among SS whereby rates will be differentially lower in the intervention arm as compared to standard care for Black and Hispanic SS's relative to non-Black or non-Hispanic SS's. Adjusting for stratification variables, we will use generalized linear modeling to evaluate the rate of acute care visits in SS (as measured by number of ED visits + urgent care visits) by 12 months follow-up as a function of treatment, Black or Hispanic identification, and the interaction of treatment and Black or Hispanic identification.

Race

Regarding the analysis of the hypothesized differential effects of treatment on number of acute care visits at twelve months as a function of race, for each individual i :

$$\ln(\lambda_{i \text{ at 12 Months}}) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Race}_i + \beta_6 \times \text{Condition}_i + \beta_7 \times \text{Condition}_i \times \text{Race}_i$$

Where $\text{Number of Acute Care Visits at 12 Months}_i \sim \text{Poisson}(\lambda_{i \text{ at 12 Months}})$ with offset term for log of the exposure amount.

If the model fails to meet Poisson assumptions, analyses will use alternative strategies including application of scaling coefficients or robust estimators, evaluation of the fit of negative binomial as well as zero-inflated and truncated.

Hispanic Ethnicity

Regarding the analysis of the hypothesized differential effects of treatment on number of acute care visits at twelve months, for each individual i :

$$\ln(\lambda_{i \text{ at 12 Months}}) = \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Hispanic Ethnicity}_i + \beta_6 \times \text{Condition}_i + \beta_7 \times \text{Condition}_i \times \text{Hispanic Ethnicity}_i$$

Where $\text{Number of Acute Care Visits at 12 Months}_i \sim \text{Poisson}(\lambda_{i \text{ at 12 Months}})$ and offset term for log of the exposure amount.

In modeling the number of acute care visits, if the model fails to meet Poisson assumptions, analyses will use alternative strategies including application of scaling coefficients or robust estimators, evaluation of the fit of negative binomial as well as zero-inflated and truncated distributions as indicated by the data.

Hypothesis 2.c.1: At 12 months, the intervention will demonstrate differential effects on the probability of readmission among SS whereby probability of readmission will be differentially lower in the intervention arm as compared to standard care for Black and Hispanic SS's relative to non-Black or non-Hispanic SS's. We will use logistic regression analysis to evaluate the probability of readmission by 12 months follow-up as a function of treatment, Black or Hispanic identification and the interaction of treatment and Black or Hispanic identification after adjusting for stratification variables.

Race

Regarding the analysis of the hypothesized differential effects of treatment on probability of readmission at three months as a function of race, for each individual i :

$$\begin{aligned} \text{Logit (Probability of Readmission}_i \text{ at 12 Months)} = & \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \\ & \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \text{Race}_i + \beta_6 \\ & \times \text{Condition}_i + \beta_7 \times \text{Condition}_i \times \text{Race}_i \end{aligned}$$

A Poisson model will also be fit with an offset term for the log of exposure amount to account for differential amount of follow-up. If the model fails to meet Poisson assumptions, analyses will use alternative strategies including application of scaling coefficients or robust estimators, evaluation of the fit of negative binomial as well as zero-inflated and truncated distributions as indicated by the data.

Hispanic Ethnicity

Regarding the analysis of the hypothesized differential effects of treatment on the probability of readmission at twelve months as a function of Hispanic ethnicity, for each individual i :

$$\begin{aligned} \text{Logit (Probability of Readmission at 12 Months}_i) = & \beta_0 + \beta_1 \times \text{Site}_i + \beta_2 \times \\ & \text{Insurance Status}_i + \beta_3 \times \text{Neuropsychological Score Strata}_i + \beta_4 \times \text{SBP}_i + \beta_5 \times \\ & \text{Hispanic Ethnicity}_i + \beta_6 \times \text{Condition}_i + \beta_7 \times \text{Condition}_i \times \text{Hispanic Ethnicity}_i \end{aligned}$$

A Poisson model will also be fit with an offset term for the log of exposure amount to account for differential amount of follow-up. If the model fails to meet Poisson assumptions, analyses will use alternative strategies including application of scaling coefficients or robust estimators, evaluation of the fit of negative binomial as well as zero-inflated and truncated distributions as indicated by the data.

6.4 ANALYSIS OF EXPLORATORY OUTCOMES AFTER TRIAL COMPLETION

Hypothesis 3.a.1: Additional measured SDOH (neighborhood social deprivation, neighborhood social vulnerability indices, food insecurity, experienced racism/discrimination, healthcare-related racism/discrimination, preferred language, and social support) will be independently associated with systolic BP at 6 months. Regularized regression analyses, with treatment as a covariate, will evaluate the independent effect of these SDOH on 6-month systolic BP. **Hypothesis 3.a.2:** Additional measured SDOH (as above) will be independently associated with composite vascular events at 12 months. Regularized logistic regression analyses, with treatment as a covariate, will evaluate the independent effect of these SDOH on probability of recurrent vascular event at 12 months. Further, we will assess for heterogeneity of treatment effect across these variables.

Both **Hypotheses 3.a.1** and **3.a.2** will utilize the same exploratory analytic strategy. Generalized linear models will be fit as in the main models, but we will precede them with variable selection using a lasso regularization, with the regularization parameter chosen using cross-validation. We will use the regularized regression to identify a parsimonious model, particularly due to the large number of SDOH variables explored.^{8,9}

7 SAMPLE SIZE JUSTIFICATION

The sample size justification is based on previous research and preliminary data on blood pressure (BP)

control post-stroke. Estimates from prior studies suggest that approximately 50% of stroke survivors (SS) have uncontrolled BP at 3 to 6 months post-stroke. To ensure an adequate sample size for the primary outcome, a conservative estimate of 50% BP control in the standard care group is used. It is hypothesized that there will be at least a 15% difference between groups at 6 months (65% in the VIRTUAL arm). With 90% power and alpha set at 0.05, a total of 454 SS will be retained in the study. Accounting for an estimated attrition rate of 15%, 534 participants will be enrolled over 36 months. While the sample size estimates are based on frequentist statistics, both frequentist and Bayesian approaches will be used for analysis.

8 SAFETY ANALYSES

8.1 OVERVIEW

The CCCT will produce SMC reports semi-annually and at the request of the SMC. The Closed Session SMC reports will show tables by partially blinded treatment group (Groups A and B).

8.2 DEFINITION OF ADVERSE EVENTS

All adverse events will be summarized by “preferred term” and associated system-organ class according to the MedDRA dictionary and by treatment group in terms of frequency of the event and number of subjects having the event. For the SMC closed report, severity and relatedness to the intervention will also be shown by treatment group.

8.2.1 Un-Blinded Statistical Monitoring

For the main safety endpoints, in each semi-annual SMC report, the proportions by treatment group and corresponding 99% confidence intervals will be provided. The rate ratio and corresponding 99% confidence intervals comparing each experimental arm to the standard arm will also be provided.

Tight control over type I error inflations related to interim safety analysis is not required; however, given the semi-annual safety review, a confidence level of 99% is proposed in order to nominally protect against false positive safety declarations. This proposed monitoring approach may be modified prior to study start in consultation with the SMC, who may wish to see the data presented more frequently or with alternative confidence levels.

At the conclusion of the study, the cumulative incidence of safety events will be compared using generalized linear modeling as a function of treatment after adjusting for stratification variables. AE's of particular interest include falls, syncope, and impaired renal function.

9 REFERENCES

1. Enders CK. Applied missing data analysis. New York; London: Guilford, 2010.
2. Fitzmaurice GM, Laird NM. Generalized linear mixture models for handling nonignorable dropouts in longitudinal studies. *Biostatistics* 2000;1(2): 141-156.
3. Al Jerdi S, Akhtar N, Mahfoud Z, Kamran S, Shuaib A. Major cardiovascular events in patients presenting with acute stroke: a 5-year follow-up study in patients who had ischaemic stroke and stroke mimics. *BMJ Open* 2022;12(3):e053059. DOI: 10.1136/bmjopen-2021-053059.
4. Cinciripini PM, Green CE, Robinson JD, et al. Benefits of varenicline vs. bupropion for smoking cessation: a Bayesian analysis of the interaction of reward sensitivity and treatment. *Psychopharmacology* 2017;234(11):1769-1779. (In eng). DOI: 10.1007/s00213-017-4580-2.
5. Dixon DO, Simon R. Bayesian subset analysis. *Biometrics* 1991;47(3):871-881. (In eng).

6. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000;355(9209):1064-1069. (In eng). DOI: 10.1016/S0140-6736(00)02039-0.
7. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statistics in medicine* 2002;21(19):2917-2930. (In eng). DOI: 10.1002/sim.1296.
8. Kuhn M, Johnson K. Applied predictive modeling. New York, NY: Springer, 2013.
9. Hastie TTR, Friedman J. The Elements of Statistical Learning. Second Edition. New York, NY: Springer, 2009.