ACTG A5395

Primary Statistical Analysis Plan

Version 2.0

A Randomized, Placebo-Controlled, Double-Blind, Trial to Evaluate the Efficacy of Hydroxychloroquine and Azithromycin to Prevent Hospitalization or Death in Persons with COVID-19

Protocol Version 2.0 with CM #1

ClinicalTrials.gov Identifier: NCT04358068

May 13, 2020

This is ACTG A5395 SAP Version 2.0 with names of authors, names of publication writing team members and analysis timeline redacted. Due to the early termination of the study, protocol objectives 1.3.2, 1.3.4, 1.4.3 and 1.4.4 were not pursued.

Table of Contents

Version History

Version	Changes Made	Date Finalized
1	Original Version	05/08/2020
2	SAP update implemented prior to enrollment of any participants. Updates include: corrected typos and approaches to handle missing values when analyzing self-reported symptoms.	5/13/2020

1   Introduction

1.1   Purpose

This Primary Statistical Analysis Plan (SAP) describes the proposed content for the primary statistical analysis report of ACTG A5395, which addresses the primary, secondary and a subset of exploratory objectives of the study. This document also describes the primary and secondary outcome measures for which results will be posted on ClinicalTrials.gov and that will be included in the primary manuscript. The Primary SAP outlines the general statistical approaches that will be used in the analysis of the study and has been developed to facilitate discussion of the statistical analysis components among the study team; and to provide agreement between the study team and statisticians regarding the statistical analyses to be performed and presented in the Primary Analysis Report.

Analyses for the Primary Analysis Report will be finalized once the last participant has completed the Day 20 study visit, all queries on data through Day 20 have been resolved, and the study data has been frozen; details on the analysis timeline are located in Appendix 3. In addition, all primary and secondary outcomes outlined in this SAP will be submitted to ClinicalTrials.gov within 1 year of the Primary Completion Date (PCD); the PCD is defined as the day of the last "Day 20" evaluation among all study participants.

Additional specifications including detailed outlines of tables, figures, and coding descriptions are included in the Analysis Implementation Plan (AIP).

1.2   Key Updates to the SAP

The study will open under protocol version 2.0 with Clarification Memo #1; therefore, the first statistical analysis plan (version 1.0) is based on this protocol version.

2   Study Overview

2.1   Study Design

A5395 is a phase IIB, double-blind, placebo controlled, randomized trial designed to compare the efficacy of hydroxychloroquine (HCQ) plus azithromycin (Azithro) versus placebo to prevent hospitalization and death in symptomatic adult outpatients with COVID-19.

Eligible participants will receive study treatment for 7 days followed by 23 weeks of follow up. During this time, participants will be assessed for the primary outcome through Day 20, and will then have remote visits at week 12 and 24 for assessments of long-term outcomes.

The study population consists of adults (≥ 18 years of age) with documented SARS-CoV-2 infection who are experiencing at least one of fever, cough, or shortness of breath, but do not require hospitalization.

2.2   Hypothesis

Hydroxychloroquine (HCQ) and Azithromycin (Azithro) will prevent hospitalization and death in persons with symptomatic SARS-CoV-2 infection.

2.3   Study Objectives

This Primary SAP addresses the following primary, secondary, and exploratory objectives listed in the study protocol. Other exploratory study objectives in the protocol will be addressed in subsequent analysis plans.

2.3.1   Primary Objectives

To determine if HCQ and Azithro will prevent the composite endpoint of either hospitalization or death by 21 days after study entry. Hospitalization is defined as ≥24 hours of acute care [Protocol Objective 1.2.1].

2.3.2   Secondary Objectives

1) To determine safety of HCQ and Azithro including 1) adverse events (AEs) leading to early discontinuation of treatment with HCQ and Azithro or 2) cardiac AEs [Protocol Objective 1.3.1].

2) To determine if HCQ and Azithro reduce the frequency of detection and levels of SARS-CoV-2 RNA in site-collected collected nasopharyngeal (NP) swabs and self-collected swabs in subset of participants [Protocol Objective 1.3.2].

3) To determine if HCQ and Azithro change the severity and duration of self-reported symptom experience of COVID-19 [Protocol Objective 1.3.3].

4) To determine if HCQ and Azithro will prevent the composite endpoint of either hospitalization or death by 24 weeks after study entry. Hospitalization is defined as ≥24 hours of acute care [Protocol Objective 1.3.4; this objective involves longer-term follow-up and will be addressed as a supplement to the Primary Analysis Report]

2.3.3   Exploratory Objectives

1) To explore differences in outcomes between HCQ and Azithro versus placebo treatment groups among subgroups of the population, notably by sex, race/ethnicity, and risk groups defined by age and comorbidities [Protocol Objective 1.4.3].

2) To explore possible predictors of outcomes across the study population, notably sex, race/ethnicity, and risk groups defined by age and co-morbidities [Protocol Objective 1.4.4].

The other exploratory objectives in the protocol will be addressed in separate SAP(s).

2.4   Overview of Sample Size Considerations

The proposed sample size is 2000 participants who take the first (confirmed) dose of study treatment (approximately 1000 per arm). Participants who are randomized but do not take the first (confirmed) dose of study treatment will not be followed and will be replaced. This sample size has been chosen to provide 90% power to detect a relative reduction of 33.3% in the proportion of participants hospitalized/dying between the study arms (HCQ and Azithro vs placebo). This is based on the following assumptions:

- Proportion hospitalized/dying in the placebo arm is 15%;
- Two-sided test of two proportions with 5% Type I error rate;
- Three interim analyses and one final analysis, equally spaced, with stopping guideline determined using the Lan-DeMets spending function approach with an O'Brien and Fleming boundary;
- Allowance for 5% of participants to be lost-to-follow-up prior to being hospitalized or dying.

Further details on the assumptions and sample size calculation are provided in protocol section 10.4.

2.5   Overview of Formal Interim Monitoring

A NIAID-appointed Data and Safety Monitoring Board (DSMB) will undertake reviews of interim data from the study to help ensure the safety of participants in the study, and to recommend changes to the study including termination or modification for safety reasons or if there is persuasive evidence of efficacy or lack of efficacy of HCQ and Azithro versus placebo in preventing hospitalizations and deaths. At each interim review, the DSMB will review summaries of data by unblinded randomized treatment arm for the primary outcome of hospitalization/death, losses to follow-up, adverse events, self-report of treatment cross-over, and the secondary outcomes of death, study treatment discontinuations (and associated reasons), duration of symptoms of COVID-19 and duration of fever. By-stratum summaries will also be reviewed.

For monitoring the primary efficacy outcome, the O'Brien Fleming boundary will be used as the stopping guideline, implemented using the Lan-DeMets spending function. Because at early interim reviews the O'Brien Fleming boundary requires extreme p-values to demonstrate greater efficacy of the treatment arm compared to placebo, a nominal two-sided p-value of < 0.001 will be used.

There will be a review of safety data by the DSMB when 250 participants have completed Day 20 of follow-up. Unless otherwise recommended by the DSMB, interim efficacy analyses will occur weekly from when approximately 500 participants have been followed for the primary outcome assessed at Day 20. If the total number of hospitalizations or deaths is higher than anticipated, the first interim efficacy review may occur earlier than planned. This would occur when approximately 62 participants in the two arms combined have been hospitalized or have died (62 is the expected number based on 15% hospitalizations/deaths in the placebo arm and 10% in the HCQ and Azithro arm, with an interim sample size of 500 participants).

In considering possible modifications to the study or termination of the study, the DSMB may consider other outcome measures besides the primary composite hospitalization/death outcome measure, or differences within subgroups. There is the possibility that differences between the treatment arms may be observed at an early study time point (for example cumulative proportion at Day 6); however, the overall goal of the study is to prevent hospitalization and deaths regardless of the timing, and therefore the focus of the treatment arm comparisons will be at Day 20.

The team does not recommend the DSMB monitor for statistical futility (i.e., stopping early for the absence of difference between arms). The DSMB will monitor operational futility. With respect to operational futility, the DSMB may recommend modification or termination of the study if the proportion hospitalized/dying in the control arm is low. In addition, the DSMB will monitor the loss to follow-up (LTFU) rate. As a benchmark, an overall LTFU rate of more than 10% would be cause for concern. The DSMB should also monitor the treatment crossover in the placebo arm by participant self-report and, if available, by PK levels of HCQ and Azithro. As a benchmark, a crossover rate of more than 10% in the placebo arm would be a cause for concern.

Additional details on interim monitoring are provided in protocol section 10.5.

3   Outcome Measures

All outcome measures are copied from the protocol. Additions, notes, and modifications to outcomes are shown in blue.

3.1   Primary Outcome Measures

1) Death from any cause or hospitalization during the 21-day period from and including the day of the first (confirmed) dose of study treatment. [For Primary Objective]

   Hospitalization is defined as requiring at least 24 hours of acute care in a hospital or similar acute care facility, including Emergency Rooms or temporary facilities instituted to address needs during the COVID-19 pandemic. Evaluation at a hospital or similar facility with less than 24 hours of acute care is not considered a hospitalization.

   Note: This will include a 24-hour hospitalization that is initiated on Day 20.

3.2   Secondary Outcome Measures

1) Death from any cause during the 21-day period from and including the day of the first (confirmed) dose of study treatment. [Supplementary information for Primary Objective]

2) Death from any cause, or hospitalization, or any urgent visit to an emergency room or clinic during the 21-day period from and including the day of the first (confirmed) dose of study treatment. [Supplementary information for Primary Objective]

   Note: This outcome includes instances where participants required acute care for less than 24 hours.

3) Death from any cause or hospitalization during the 24-week period from and including the day of the first (confirmed) dose of study treatment. [For Secondary Objective 4, with follow-up beyond the PCD]

   Note: The definition of hospitalization is the same as the primary outcome.

4) Premature discontinuation of study treatment due to an adverse event. [For Secondary Objective 1]

   Note: Premature discontinuation of study treatment is defined as a permanent discontinuation of either study treatment (HCQ/Placebo and/or Azithro/Placebo).

5) Occurrence of any cardiac adverse events from start of study treatment through Day 20. [For Secondary Objective 1]

   Note: See Appendix 1 for list of applicable MedDRA codes.

6) Any occurrence of fainting by self-report from participant's daily diary cards from the time of starting treatment through Day 20. [For Secondary Objective 1]

   Fainting is collected as part of a set of targeted symptoms collected on participant daily diary cards, which are completed prior to starting treatment and then in the evening of each day from Day 0 to Day 20. Fainting is recorded as absent (score 0), mild (1), moderate (2), or severe (3); scores of > 0 are defined as occurrence of fainting.

7) Duration of fever defined as the time from start of study treatment to the last day in the participant's daily diary card on which a temperature greater than 100.4^0 F was recorded or a potentially anti-pyretic drug, such as acetaminophen or ibuprofen, was taken. Diary cards are completed at entry prior to starting treatment and then in the evening of each day from Day 0 to Day 20. [For Secondary Objective 3]

8) Duration of symptoms associated with COVID-19 defined as the time from start of study treatment to the last day in the participant's daily diary card on which a moderate or worse targeted symptom was recorded. Diary cards are completed at entry prior to starting treatment and then in the evening of each day from Day 0 to Day 20. [For Secondary Objective 3]

The set of target symptoms are the same ones that are used in many influenza studies with the addition of the symptoms of loss of smell and loss of taste, which have been associated with COVID-19. The ones from influenza studies are: cough, shortness of breath, feeling feverish, fatigue, muscle aches, diarrhea, vomiting, nausea, headache, sore throat, nasal obstruction (stuffy nose), nasal discharge (runny nose). These symptoms are listed individually on the study diary that each participant will complete each day. The scoring is that used in influenza studies whereby participants grade each symptom as absent (score 0), mild (1), moderate (2), or severe (3).

9) Participant-specific area under the curve (AUC) of the total symptom score associated with COVID-19 over time (through Day 20) defined as the sum of scores for the targeted symptoms (defined above) in the participant's daily diary record (each individual symptom is scored from 0 to 3). Diary cards are completed at entry prior to starting treatment and then in the evening of each day from Day 0 to Day 20. [For Secondary Objective 3]

Note: This outcome measure can be considered as a composite severity ordinal outcome, as information on hospitalization and death is incorporated with the symptom-score-based AUCs for participants who did not die and were not hospitalized.

10) Time to self-reported return to (pre-COVID) usual health defined as the time from start of study treatment to the first day in the participant's daily diary card on which they reported *Yes* with no subsequent reports of *No*. Diary cards are completed at entry prior to starting treatment and then in the evening of each day from Day 0 to Day 20. [For Secondary Objective 3]

11) SARS-CoV-2 RNA level (detectable versus not detectable) from self-collected nasal swabs at entry and days 6 and 20 among a subset of participants. [For Secondary Objective 2]

12) SARS-CoV-2 RNA level (detectable versus not detectable) from site-collected NP swabs at entry and days 6 and 20 among a subset of participants. [For Secondary Objective 2]

13) SARS-CoV-2 RNA level (continuous, reported as Cycle Number) from self-collected nasal swabs at entry and days 6 and 20 among a subset of participants. [For Secondary Objective 2]

14) SARS-CoV-2 RNA level (continuous, reported as Cycle Number) from site-collected NP swabs at entry and days 6 and 20 among a subset of participants. [For Secondary Objective 2]

4   Statistical Principles

4.1   General Considerations

All analyses will include all randomized participants who received the first (confirmed) dose of study treatment according to a modified intent-to-treat approach (mITT), with the exception of analyses to address protocol objective 1.3.2, which will further be restricted to those in the subset who provided NP and/or nasal swabs. Participants who enroll, but do not receive their first (confirmed) dose of study treatment will not be included in any analyses.

Study visit windows for reporting are as defined in the protocol. Key study visits are Entry (Day 0), Day 6 or 7, Day 20, and Week 24. Baseline is defined as the last available measure prior to the initiation of study treatment.

Entry (Day 0):   First dose of study treatment occurs.

Day 6 or 7:   Last day on study treatment.

Note: Depending on the timing of the first dose on Day 0 the last dose of study treatment may occur at the end of Day 6 or the beginning of Day 7

Day 20:   Last day primary outcome may occur.

Week 24:   Last study visit.

Statistical comparison across treatment arms of baseline characteristics are not planned because the study is randomized and placebo-controlled; hence, any differences should reflect chance variation.

Analyses of primary and secondary outcomes will not adjust for multiple comparisons. Analyses of primary outcomes will adjust for the multiple interim reviews using group sequential methods.

Continuous variables will be summarized using median, interquartile range (Q1 and Q3), 10th and 90th percentile, and min and max; categorical variables will be summarized using frequency and percentage.

NIH requires the primary analyses of treatment comparisons to be summarized by sex/gender and by race/ethnicity and that treatment interactions with sex/gender and race/ethnicity be evaluated.

4.2   Analysis Approaches

4.2.1   Analysis of the Primary Objective

Analysis Population

The analyses of the primary objective will include all randomized participants who received the first (confirmed) dose of study treatment according to a modified intent-to-treat (mITT) approach. Participants who start study treatment outside of the protocol-defined study windows (i.e. more than 96 hours from SARS-CoV-2 positive test) will be included; however, participants who enroll, but do not receive their first dose of study treatment will not be included in the analyses.

Note: Participants who are randomized but do not take the first (confirmed) dose of study treatment are not to be followed and will be replaced.

Analysis Methods

The analysis of the primary outcome will compare the cumulative proportion of participants hospitalized or dying (from any cause), from Day 0 through to Day 20, between randomized arms using a ratio of proportions. For analysis purposes, the integer scale will be used as the time scale, where Study Day 0 is considered day 1 and Study Day 20 is considered day 21. The cumulative proportion will be estimated for each randomized arm in each risk stratum (based on stratification factor 'high' vs 'low' risk) using Kaplan-Meier methods to account for losses to follow up (and differential follow-up at the time interim reviews). Participants will have follow up censored at the date they were last known to be alive and not hospitalized through Day 20 (i.e. integer day 21). The primary analysis assumes non-informative censoring.

Within each stratum, the absolute difference in the estimated log-cumulative proportion will be calculated between randomized arms; the variance for this difference will be obtained using Greenwood's formula. The stratum-specific estimated differences will then be combined across strata, weighted by the inverse of the stratum-specific variance of the estimated difference, to obtain an overall estimate. Two-sided 95% confidence intervals (CIs) and p-value (for the test of no difference between arms) will be obtained, which adjust for the interim analyses; a nominal 95% CI and p-value will also be provided. By-stratum estimates and corresponding nominal 95% CIs and nominal p-values will also be provided, and will not be adjusted for interim looks.

Sensitivity Analyses of the Primary Outcome

The following sensitivity analyses are included to evaluate impact of different assumptions on the inference of the primary treatment comparisons.

1)   Evaluate the composite outcome of being hospitalized, dying, or lost to follow up.

Approach: Repeat the primary analysis, but assume all participants who prematurely discontinue the study prior to Day 20, who are unable to be contacted by the site to ascertain outcomes after discontinuation, had a primary event at Day 20.

2) Evaluate the impact of participants who delay starting study treatment (more than 96 hours after positive SARS-CoV-2 test, per protocol).

   Approach: Repeat primary analysis, but exclude participants who started study treatment more than 96 hours after the positive SARS-CoV-2 test, as defined in the protocol.

3) Evaluate the impact of differential loss-to-follow-up.

   Approach: In the event that differences are observed between the primary analysis and the first sensitivity analysis, the impact of participants being LTFU will be explored using inverse probability of censoring weights (IPCW). The primary analysis will be repeated but, within each arm, participants who are not loss to follow-up will be weighted using inverse probability of censoring weights (IPCW) determined by baseline variables that predict loss to follow up.

4) Evaluate the impact of participants enrolling from the same household.

   Approach: Repeat the primary analysis only including the first participant who enrolled from each household.

   In the event that differences are observed between the primary analysis and this sensitivity analysis, analysis methods that account for clustering will be considered, if feasible.

Subgroup Analyses of the Primary Outcome

To evaluate the effect of treatment in specific populations, the primary outcome will be assessed among different subgroups. The same approaches outlined for the primary analysis will be implemented for each subgroup. The difference between treatment arms in the log-proportion will be estimated within subgroup, and compared between subgroups by determining the difference of the treatment differences. In the event that the proportion of participants in each subgroup is low, or the number of events are low, descriptive summaries of the number of hospitalizations and deaths will be done. Pre-specified subgroups of interest include:

1) Sex (Male sex at birth, female sex at birth)
2) Race (white, non-white)
3) Ethnicity (Hispanic, non-Hispanic)
4) 'Risk of Severe Disease' Strata (<60 years and no comorbidities, ≥ 60 years or at least one comorbidity)
5) Age Group (<60, ≥60)
6) Co-morbidity Status (no comorbidities, at least one comorbidity)
7) Calendar days from first symptom associated with COVID-19 to start of study treatment (≤ median days, >median days)

<u>Supportive Analyses of the Primary Outcome</u>

Secondary outcomes 1 and 2 are included as supportive to the primary outcome. The cumulative proportion of participants dying (from any cause) from Day 0 through to Day 20 (i.e. integer day 1 through 21), and the cumulative proportion of participants dying (from any cause), hospitalized, or who had an urgent visit to an emergency room or clinic from Day 0 through to Day 20 (i.e. integer day 1 through 21), will be analyzed in the same manner as the primary outcome.

Secondary outcome 3, which corresponds to secondary objective 1.3.4 from the protocol, will be analyzed as part a supplemental report (for 24 weeks of follow-up) in the same manner as the primary outcome. This outcome evaluates hospitalizations and deaths through week 24, and will censor participant follow-up at the date they were last known to be alive and not hospitalized through Study Day 167, which corresponds to integer day 168.

4.2.2   Analyses of Secondary Safety Objectives

Premature discontinuation of study treatment due to an adverse event (tolerability), any occurrence of fainting by self-report from participant's daily diaries from start of study treatment through Day 20, and occurrence of any cardiac adverse events from start of study treatment through Day 20 will be analyzed in the following manner. For each outcome, the proportion of participants who experienced each outcome will be estimated and compared between randomized arms using log-binomial regression, with log link and adjustment for a participant's risk stratum, in order to obtain a risk ratio estimate. In the event the log-binomial regression model fails to converge, a Poisson regression model with robust variance and log-link will be used instead.

The cardiac AEs included in this analysis are shown in Appendix 1 by MedDRA codes; these were chosen a priori by the study chairs.

4.2.3   Analyses of Secondary Efficacy Objectives

<u>Analyses of Symptom and Fever Duration</u>

Analyses of duration of fever, duration of symptoms, and time to self-reported return to usual health will be restricted to those who received first (confirmed) dose of study treatment (i.e., mITT population). However, duration of fever analyses will further restrict to those participants who ever reported a temperature reading from a thermometer.

Duration of fever, duration of symptoms, and time to self-reported return to usual health will be summarized with descriptive statistics. Participant specific durations for each outcome will be compared between randomized arms using a two-sided Wilcoxon test stratified by risk stratum with a two-sided 5% type I error rate.

Participants who never report fever or use of anti-pyretic medications will be assigned a duration of fever of zero days. A sensitivity analysis will define duration of fever as the time from start of study treatment to the last day in the participant's daily diary card on which a temperature greater than 100.4^0 F was recorded, and will not make any special considerations for participants who

indicated use of potentially anti-pyretic drugs (i.e. will not include the use of a potentially anti-pyretic drug in the definition). In this sensitivity analysis, those who never report fever will be assigned duration of fever of zero days.

Participants who have missing diary cards due to death will be ranked as the worst outcome in these analyses. Participants who have missing diary records due to hospitalization will have:

- Daily symptoms assumed to be *at least moderate* during hospitalization;
- Self-reported return to usual health assumed to be *No* during hospitalization;
- Fever assumed to be *Yes* during hospitalization.

Participants who have missing diary cards not due to death or hospitalization will have all available data included in analysis and missing data will be ignored in the analyses of self-report return to usual health and fever duration; missing data in analyses of daily symptom will be interpolated based on the preceding and succeeding symptoms.

Supplementary analyses will analyze duration of fever and symptoms using Kaplan-Meier methods. These analyses will compare the time to first day of two successive assessments without fever and time to first day of two successive assessments without moderate or severe symptoms between study arms using the Wilcoxon test stratified by risk stratum. These analyses will be restricted to participants who report fever or symptoms, respectively, at Day 0 on the pre-treatment diary cards. Analyses using IPCW may be considered to account for LTFU.

Analyses of Total Symptom Score

Total symptom score will be calculated for each daily diary card for each person by summing the individual symptoms reported on the daily diary; symptom scores range from zero to three (0=absent, 1=mild, 2=moderate, 3=severe). Participant-specific areas under the curve (AUC) over time will be calculated using the trapezoidal rule and is defined as the area below the line formed by joining total symptom scores on each daily diary card from the pre-treatment score on Day 0 through to Day 20. For analysis purposes, pre-treatment diary cards on Day 0 are assigned integer day 0; all post-treatment diary cards assigned integer days in the same manner as the primary analysis. The AUCs will be rescaled by time by dividing by 21, corresponding to (number of daily diary cards between pre-treatment Day 0 and Day 20), in order to provide results on a symptom scale from 0 to 42. Participant-specific AUCs will be compared between randomized arms using a two-sided Wilcoxon test stratified by risk stratum with a 5% type I error rate.

Participants who are hospitalized or die between Day 0 and Day 20 will not have their AUC calculated. Participants who die will be ranked as having the worst AUC outcome in the analysis. All deaths will be ranked as the worst outcome regardless of when they occur between Day 0 and Day 20. Participants who are hospitalized will be ranked as having the next worse outcome (compared to death) in the analysis, with ranking determined based on the duration of hospitalization between Day 0 and Day 20; participants with the longest duration will be ranked as having worse outcomes than those with shorter durations. Duration is defined as the number of days hospitalized that includes the day of admission and the day of discharge (i.e., possible range from 1 to 21 days) between Day 0 and Day 20.

Participants who have incomplete diary cards for reasons other than hospitalization or death will be handled in the following manner:

1) Participants who are missing pre-treatment symptom scores will have their scores imputed as the mean pre-treatment total score among participants who have the same total symptom scores reported on the post-treatment Day 0 diary cards;
2) Participants who have intermittent days with no symptom scores reported (i.e., all scores missing), the daily diary cards with missing scores will be ignored in AUC calculation, which is analogous to interpolating the total symptom scores;
3) Participants who have diary cards with some, but not all symptoms scores reported, the missing symptoms scores will be interpolated based on the preceding and succeeding available scores for a given symptom;
4) Participants who stop completing their symptom diaries before Day 20 will have their last total symptom score carried forward through Day 20, and their AUC calculation done as noted above. Methods such as multiple imputation or IPCW may be considered if more than 10% of participants in either arm stop completing their diaries before Day 20 for reasons other than death or hospitalization.

Analyses of SARS-CoV-2 RNA

These analyses will be conducted among the subset of participants who had nasal and/or NP swabs collected, respectively, and will include all participants who received first (confirmed) dose of study treatment. Sensitivity analysis will further restrict to those with detectable SARS-CoV-2 RNA at Day 0. Analysis will be conducted separately for each swab type. SARS-CoV-2 RNA is quantified by the target cycle number (CN), which is the point in the PCR at which SARS-CoV-2 virus is detected; the number of cycles is inversely related to the amount of virus. Assays will be run in batch after the primary completion date.

The proportion of participants with detectable SARS-CoV-2 RNA will be compared between randomized arms using log-binominal regression with log-link and adjustment for risk stratum, pre-treatment (Day 0) SARS-CoV-2 CN. For each time point (Day 6 or Day 20), the model will include main effects for treatment arm and time, and an interaction between time and treatment to evaluate differences between arms. In addition, a joint test of treatment across the two time points (2 degrees of freedom) will also be assessed. This model will be fitted using generalized estimating equations (GEE) to handle the repeated measurements with unstructured covariance matrix. With this model, the comparison between randomized arms will use a two-sided Wald-test with 5% type I error rate.

Non-parametric Wilcoxon rank-sum tests stratified by risk stratum with a 5% type I error rate will compare SARS-CoV-2 RNA cycle numbers (continuous) between randomized arms, separately at Day 6 and Day 20; results without cycle numbers due to target not detected will be imputed as the highest rank. A composite test, simultaneously analyzing Day 6 and Day 20 data will also be performed (DeLong ER, DeLong DM, Clarke-Pearson DL. Biometrics. 1988 Sep 1:837-45).

Missing data are assumed to be missing completely at random (MCAR) and will be ignored in analysis. Sensitivity analyses will address possible informative missingness. In particular, results

that are missing due to participant hospitalization or death will be imputed as *detectable* in the dichotomous analysis and as the lowest rank the continuous analysis.

4.3 Interim Analyses

The Therapeutics and Prevention DSMB will review interim data from the study including descriptive summaries of study conduct and adverse events, and efficacy analyses that contrast treatment arms. The primary outcome of death or hospitalization will be compared between treatment arms using the statistical methods outlined in this SAP; the secondary outcome of death from any cause will be also be compared between treatments arms. Interim efficacy analyses are planned to occur weekly from when approximately 500 participants have been followed for the primary outcome assessed at Day 20, or earlier if the total number of hospitalizations or deaths is higher than anticipated (See SAP Section 2.5 or Protocol Section 10.5).

At each interim review, the stopping boundary for the primary analysis will be determined based on the proportion of planned maximum information that is available at the given review. The statistical information (Fisher's Information) at a given review will be calculated using the inverse of the variance (square of standard error) obtained from Greenwood's formula as part of the primary analysis. The maximum information will be pre-determined using the following formula (Tsiatis AA. Statistics in medicine. 2006 Oct 15;25(19):3236-44):

$$MI = \left\{ \frac{\left(Z_{\alpha/2} + Z_\beta\right)^2}{\delta_A^2} \right\} * (Inflation\ Factor) = \frac{(1.96 + 1.28)^2}{\left\{ \ln\left(\frac{0.10}{0.15}\right) \right\}^2} * 1.03 = 65.8.$$

However, at early interim reviews, where the O'Brien Fleming spending function is particularly extreme, a nominal two-sided p-value of <0.001 will be used instead.

The interim efficacy analyses will also include two sensitivity analyses. The first will evaluate the potential impact of delayed ascertainment of the primary endpoint, as requested by the DSMB and using an approach suggested by a DSMB statistician (personal correspondence A.A. Tsiatis). The second will assume all participants who prematurely discontinue the study prior to Day 20, who are unable to be contacted by the site to ascertain outcomes after discontinuation, had a primary event at Day 20.

5 Report Contents

Note: Tables in the primary analysis report will summarize outcomes overall, by treatment arm, and by risk strata as appropriate.

5.1 Flow Diagram

- CONSORT (Flow) diagram including:
    - Number of individuals screened;
    - Number of participants enrolled/randomized and reasons why participants did not enroll;

- o Number of participants who started study treatment and reasons why participants did not start treatment;
- o Number of participants who discontinued study treatment and study prematurely with reasons why.

5.2   Accrual and Eligibility

- Summary of number participants enrolled by week and site;
- Summary of number participants who started treatment by week and site;
- Summary of participants with eligibility violations.

5.3   Baseline Characteristics

- Summary of age (including <60 vs ≥ 60), sex, gender identity, race, ethnicity;
- Summary of risk stratification (stratification factor "high" vs "low");
- Summary of comorbidities (defined in protocol section 6.3.4);
- Summary of smoking status including cigarettes, cigars, vaping, and marijuana;
- Summary of pregnancies;
- Summary of COVD-19 symptoms:
  - o Duration of symptoms prior to start of study treatment
  - o First symptoms (self-report)
  - o Symptoms at screening: fever, cough and/or shortness of breath;
- Summary of number of participants from same household;
- Summary of symptom scores from participant diaries collected prior to starting treatment;
- Summary of concomitant medication(s) used at entry obtained from participant diaries; collected prior to starting treatment.

5.4   Study Status

- Summary of participants on vs. off study;
- Summary of premature study discontinuations with reasons;
- Summary of follow-up time (Day 0 to last contact).

5.5   Study Treatment Status

- Summary of reasons for not starting study treatment;
- Summary of days from randomization to first dose of study treatment;
- Summary of participants on vs. off study treatment;
- Summary of premature treatment discontinuations with reasons:
  - o Secondary Safety [Protocol Objective 1.3.1]: Proportion of participants who discontinued study treatment due to AE will be compared between arms as noted in SAP section 4.2.2;
- Summary of treatment adherence including:
  - o Summary of number of missed doses of HCQ and Azithro overall and by study day with reasons;

- Summary of self-reported use of HCQ and/or Azithro.

5.6 Adverse Events

- Summary of all new, post-entry reportable AEs by MedDRA code:
    - All grade 3 or higher AEs
    - All cardiac AEs regardless of grade
    - All AEs that led to a change in study treatment regardless of grade
    - All AEs meeting SAE definition or EAE reporting requirement;
- Summary of cases of torsade de pointes (TdP), all ventricular arrhythmias, and sudden cardiac deaths:
    - Defined as events with higher level group terms (HLGTs): 10000032, 10037908, 10047283, or deaths with primary cause of death field including *sudden death* on the eCRF;
- Secondary Safety [Protocol Objective 1.3.1]: Proportion of participants who experienced a cardiac AE (defined in Appendix 1) and proportion of participants who self-report fainting will be compared between arms as noted in SAP section 4.2.2.

5.7 Pregnancy

- Summary of individuals who are pregnant at entry and who become pregnant while on study, including days/weeks from Day 0, risk stratum, pregnancy outcomes, and maternal adverse events.

5.8 Mortality

- Summary of number of deaths, primary cause of death, time from Day 0 to death, risk stratum, and demographics;
- Secondary Efficacy [Supplementary for Primary Objective]: Proportion of participants dying from any cause will be compared between treatment arms as noted in SAP section 4.2.1.

5.9 Efficacy

- Primary Efficacy [Protocol Objective 1.2.1]: Proportion of participants hospitalized or dying from any cause will be compared between treatment arms as noted in SAP section 4.2.1. This analysis includes sensitivity, supportive, and subgroup analyses;
- Secondary Efficacy [Protocol Objective 1.3.3]: Summary of duration of fever, symptoms and time to self-report return to usual health will be compared between randomized arms as noted in SAP section 4.2.3;
- Secondary Efficacy [Protocol Objective 1.3.3]: AUC of total symptom score over time will be compared between randomized arms as noted in SAP section 4.2.3.

5.10   Subset Analyses (n=200)

- Baseline Characteristics:
    - Repeat those outlined in section 5.3 among the subset
    - Summary of baseline laboratory values;
- Study Status:
    - Repeat those outlined in section 5.4;
- Study Treatment Status:
    - Summary of reasons for not starting study treatment
    - Summary of days from randomization to first dose of study treatment
    - Summary of participants on vs. off study treatment with reasons
    - Summary of treatment adherence including number of missed doses by study day, with reasons
    - Summary of self-reported use of HCQ and/or Azithro;
- Mortality and Hospitalization:
    - Summary of number of deaths, primary cause of death, time from Day 0 to death, risk stratum, and demographics
    - Summary of number of participants hospitalized;
- Symptoms:
    - Summary of duration of fever, symptoms and time to self-report return to usual health;
- Efficacy:
    - Secondary Efficacy [Protocol Objective 1.3.2]: Proportion of participants with detectable SARS-CoV-2 RNA and distribution of SARS-CoV-2 RNA levels summarized by CN will be compared between randomized arms in the subset of participants with NP and nasal swabs as noted in SAP section 4.2.3.

Appendix 1   Cardiac Adverse Events

The following Cardiac AEs were identified a priori by the study chairs to be included in the analyses of Secondary Objective 1. All AEs are described using MedDRA coding, and fall under the SOC Cardiac Disorders (SOC: 10007541).

Level	Code	Term
HLGT	10007521	Cardiac arrhythmias
HLGT	10011082	Coronary artery disorders
HLGT	10019280	Heart failures
HLGT	10028593	Myocardial disorders
HLGT	10034468	Pericardial disorders
HLGT	10046973	Cardiac valve disorders
HLGT	10082206	Cardiac disorders, signs and symptoms NEC
Among Endocardial Disorders (HLGT 10014662):		
HLT	10014660	Endocardial disorders NEC
HLT	10014686	Endocardial viral infections