STATISTICAL ANALYSIS PLAN

Efficacy of Hydroxychloroquine for Post-exposure Prophylaxis (PEP) to Prevent Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection Among Adults Exposed to Coronavirus Disease (COVID-19): A Blinded, Randomized Study

Protocol Registry Number:
NCT04328961

Effective Date: August 22, 2020

Version: 1.0

1. LIST OF ABBREVIATIONS AND ACRONYMS

A list of abbreviations used in the SAP.

Term/Abbreviation	Definition
HCQ	hydroxychloroquine
PEP	post-exposure prophylaxis
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2

2. INTRODUCTION

This is a randomized, multi-center, placebo-equivalent (ascorbic acid) controlled, blinded study of the efficacy of hydroxychloroquine (HCQ) post-exposure prophylaxis (PEP) for the prevention of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in adults exposed to the virus.

2.1. GENERAL DESIGN CONSIDERATIONS

Evaluations include safety and tolerability, SARS-CoV-2 detection, SARS-CoV-2 viral shedding, and coronavirus disease (COVID-19) diagnosis. Two study groups (one active and one receiving ascorbic acid [vitamin C], to serve as a placebo-equivalent comparator) will be enrolled to assess a daily dosing regimen administered for 14 days, which appears to be the upper limit of the incubation period of SARS-CoV-2 infection.

Up to 2000 eligible participants 18 to 80 years of age will be randomized (at the level of household) 1:1 to receive one of the following therapies:

- HCQ 400 mg orally daily for 3 days then 200 mg orally daily for an additional 11 days
- Ascorbic acid 500 orally daily for 3 days then 250 mg orally daily for 11 days

HCQ and ascorbic acid will appear similar, and taste will be partially masked as HCQ can be bitter and ascorbic acid will be sour.

During the study participants will perform the following:

- Collect mid-nasal swabs for viral detection for the primary trial endpoint
- Complete Surveys that will include questions about symptoms from both the drug regimen and virus infection, review of concomitant medications, and other pertinent topics

Note: Participants who report respiratory or other febrile illness will be referred for assessment to their primary care provider.

During the first 14 study days, participants take medication, complete Surveys, and collect mid-nasal swab to assess symptoms and virus exposure. On Day 28, a final swab is collected and Survey completed. The duration of study participation will be approximately 28 days.

2.2. Study Objectives and Endpoints

This document focuses on the statistical procedures for the primary and secondary objectives and endpoints listed in the following table.

Objectives	**Endpoints**
Primary	
To test the efficacy of HCQ (400 mg orally daily for 3 days then 200 mg orally daily for an additional 11 days, to complete 14 days) to prevent incident SARS-CoV-2 infection, compared to ascorbic acid among contacts of persons with SARS-CoV-2 infection	Polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection through Day 14
Secondary	
To determine the safety and tolerability of HCQ as SARS-CoV-2 PEP in adults	Adverse events
To test the efficacy of HCQ (400 mg orally daily for 3 days then 200 mg orally daily for an additional 11 days, to complete 14 days) to prevent incident SARS-CoV-2 infection two weeks after completing therapy, compared to ascorbic acid among contacts of persons with SARS-CoV-2 infection	Polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection through Day 28
To test the efficacy of HCQ to shorten the duration of SARS-CoV-2 shedding among those with SARS-CoV-2 infection in the HCQ PEP group	SARS-CoV-2 viral shedding by PCR
To test the efficacy of HCQ to prevent incident COVID-19	PCR-confirmed COVID-19 diagnosis post start of HCQ therapy

2.3. Randomization

Enrolled contacts were randomized 1:1 to receive either HCQ or ascorbic acid. Randomization tables were stratified by study site and contact type (household/social exposure or healthcare worker exposure) and blocked into groups of randomly varying size. The tables were programmed in R and allocation was performed via REDCap by site pharmacists or their designees. All other study staff were blinded to randomization group apart from the coordinating center's unblinded statistical team.

2.4. Sample Size and Power

Since finalization of the protocol and at the recommendation of the DSMB, the study size has been recalculated as an event-driven design. The original study was designed under 80% power with a target number of households, assuming 1 infection per household, recognizing there would

likely be more than 1 per household and therefore power would be at least 80%. In the spirit of accounting for the within household correlation without any existing estimates of what it may be, we instead assume 90% power with 50% reduction and 4 interim analyses (for 5 total) and a one-sided alpha = 0.025 and a hard futility boundary. Under these design assumptions, the design requires 93 events. Assuming a household attack rate of 10% in the placebo arm (higher than in protocol and based on emerging evidence (UM study and HH study from China that looked at masks and disinfecting), we would need to enroll 1240 participants who are negative at baseline.

3. GENERAL DATA ANALYSIS CONSIDERATIONS

3.1. ANALYSIS SET(S)

This section describes the cohorts for the primary, secondary and exploratory analyses described in this document.

Intention to Treat (ITT) Cohort: The ITT Cohort includes all contacts enrolled into the study.

Modified Intention to Treat (mITT) Cohort: Contacts who are SARS-CoV-2 negative by PCR at the baseline visit.

mITT Sensitivity Cohort: Contacts in the mITT cohort who are SARS-CoV-2 negative by PCR at the day 1-3 visits. Participants who are SARS-CoV-2 negative on two of the three days, without a SARS-CoV-2 positive result, will be included in this cohort.

Seronegative Cohort: If a validated, sensitive and specific assay is available for antibody testing, this cohort will include all mITT cohort participants who are antibody negative at baseline.

Infected Cohort: Contacts who are SARS-CoV-2 positive by PCR at the baseline visit.

PK Cohort: Participants from the DBS Sub-study with at least 1 interpretable PK sample

Household Cohort: All enrolled index cases and their household contacts. Enrolled indexes are enrolled in either the PEP Study or the companion Early Treatment Study.

Index Cohort: All enrolled index cases.

4. GENERAL ANALYSIS METHODS

Descriptive statistics that will be used to summarize continuous variables are as follows: mean and standard deviation, median and interquartile range, quartiles, range, and number of missing data values. For categorical variables, descriptive statistics that will be used include the following: frequencies, relative frequencies, and the number of missing data values. Descriptive analyses summarizing baseline and follow-up data will be stratified by site. Line, scatter and box plots will be used, as appropriate, for longitudinal data representations. A two-sided alpha level of 0.05 will be used for all statistical tests unless otherwise specified.

5. TRIAL PARTICIPANT DISPOSITION

5.1. DISPOSITION OF PARTICIPANTS

Relevant to the analyses described in this SAP, a CONSORT flow diagram will be constructed to describe how all of the participants who entered the study are accounted for including the following measures, as applicable:

The numbers of participants screened, screened out (with reasons for exclusion), enrolled, randomized, allocated to each study arm, received assigned intervention (with reasons for not receiving assigned intervention), reasons for discontinuation from follow-up (e.g., lost to follow-up, adverse event, noncompliance, etc.), and the number of participants with available data for each assay.

5.2. TREATMENT EXPOSURE

Self-reported pill taking will be reported by prescribed dosing day stratified by arm and overall. The summary measures will be number of doses reported taken and number of doses reported taken divided by number prescribed. The number and proportion of adherent participants will be reported by day and overall time.

6. BASELINE DATA

Baseline variables will include age, sex at birth, race, number of people in household, number of people in household enrolled in the study as contacts, number of infected contacts, underlying medical conditions (see below), contact type (HCW or household), household type, presence and number of children, number of bedrooms, number of bathrooms, number of days since contact with index case, duration of exposure to index case, current smoker, quarantine status hours of contact with the index case and the impact of pandemic on the participant's ability to work.

The following list of variables will be included and condensed into the following categories: metabolic disease (hypertension, diabetes mellitus type 1 and 2, and cardiovascular disease including atherosclerosis), non-communicable diseases (NCDs) (chronic cardiac disease excluding hypertension, chronic kidney disease, liver disease, chronic neurological disorder, chronic hematologic disease, rheumatologic disorder, dementia, malnutrition) infectious diseases not HIV, immunosuppressive disease (HIV, autoimmune disease and malignancy), lung disease (COPD, asthma) and other:

Chronic cardiac disease, including congenital heart disease and heart failure (not hypertension)

Hypertension

Diabetes mellitus: Type 1

Diabetes mellitus: Type 2

Chronic pulmonary disease, including COPD

Asthma (physician diagnosed)

Chronic kidney disease

Moderate or severe liver disease

Mild liver disease

Chronic neurological disorder

Malignant neoplasm

Chronic hematologic disease

HIV

Rheumatologic disorder

Dementia

Malnutrition

Autoimmune disease

7. EFFICACY/EFFECTIVENESS ANALYSES

7.1. EFFICACY ANALYSES

Final (as opposed to interim) efficacy analyses will be stratified (baseline hazard) for enrollment site and contact type and adjusted by baseline variables as noted in each section below.

7.1.1. Primary efficacy analyses

Objective: To estimate the efficacy of HCQ (400 mg orally daily for 3 days then 200 mg orally daily for an additional 11 days, to complete 14 days) to prevent incident SARS-CoV-2 infection, compared to ascorbic acid among contacts of persons with SARS-CoV-2 infection.

Outcome: Polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection through Day 14

Cohorts: mITT Cohort, Sensitivity Cohort

Definition of Endpoint: An endpoint defining PCR test is defined as one with either concordant results or Ct $\leq$ 40 and will be referred to as positive test for this analysis. PCR tests not meeting these criteria will be referred to as negative. A participant will be defined as having reached the endpoint on the midpoint between the days of collection of their first positive result and last negative result (prior to the positive result) in the first 13 days of follow-up. Participants who have confirmed COVID-19 through external evaluation in the first 13 days (e.g. hospital or clinical assessment) will be defined as having reached the endpoint at the midpoint between the last prior negative test and the day of the external diagnosis. Participants who do not reach the endpoint will have their follow-up censored at their last negative RNA test. Any uncertainty in final endpoint determination will be resolved by the blinded Endpoint Adjudication Committee.

Analysis Details:

The primary analysis will be performed on the mITT Cohort. A sensitivity analysis will be performed on the Sensitivity Cohort to account for potential missed infections at day 1.

We will use a Cox Proportional Hazards model with robust standard errors to estimate the hazard ratio and corresponding 95% confidence interval and Wald test statistic. Efficacy will be calculated as 1 minus the hazard ratio (treatment/control) from the treatment coefficient. Robust confidence intervals and p-values will be displayed. Per FDA guidance on prevention studies, the following variables will be included to increase precision: quarantine status, sex, and age of contact.

The number of evaluable participants per arm will be displayed (defined as participants with at least one PCR test result or external diagnosis) by arm with person-days of follow-up and number of endpoints per arm and prevalence of endpoints (number reached endpoint/number evaluable) with a 95% confidence interval calculated using exact binomial confidence intervals.

7.1.2. Secondary efficacy analyses

Infection by day 28

Objective: To test the efficacy of HCQ (400 mg orally daily for 3 days then 200 mg orally daily for an additional 11 days, to complete 14 days) to prevent incident SARS-CoV-2 infection two weeks after completing therapy, compared to ascorbic acid among contacts of persons with SARS-CoV-2 infection.

Outcome: Polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection through Day 28

Cohorts: mITT Cohort, Sensitivity Cohorts

Definition of Endpoint: An endpoint defining PCR test is defined as one with either concordant results or Ct $\leq$ 40 and will be referred to as positive test for this analysis. PCR tests not meeting these criteria will be referred to as negative. A participant will be defined as having reached the endpoint on the midpoint between the days of collection of their first positive and last negative PCR tests (prior to a positive test). Participants who have confirmed COVID-19 through external evaluation (e.g. hospital or clinical assessment) will be defined as having reached the endpoint at the midpoint between the last negative test (prior to external diagnosis) and the day of the external diagnosis. Participants who do not reach the endpoint will have their follow-up censored at their last negative test. Any uncertainty in final endpoint determination will be resolved by the blinded Endpoint Adjudication Committee.

Analysis Details: The same analytical approaches will be applied as for the primary efficacy analysis, including a primary analysis on the mITT Cohort and a sensitivity analysis on the Sensitivity Cohort.

Duration of shedding

Objective: To test the efficacy of HCQ to shorten the duration of SARS-CoV-2 shedding among those with SARS-CoV-2 infection.

Outcome: SARS-CoV-2 viral shedding by PCR

Cohort(s): mITT Cohort

Definition of Endpoint: For this analysis, a positive PCR test is defined as one where both targets are identified or one target is identified with Ct $\leq$ 40.

PCR tests not meeting these criteria will be referred to as negative. The endpoint will be defined as the number of days from the first positive test until virus is no longer detected (a negative test has been observed with no positive tests following). Censoring will be indicated if the participant is still shedding on day 14. Participants who have a COVID-19 related hospitalization or death will be assumed to shed through day 14 and would have their days of shedding censored as if they had been observed through day 14.

Analysis Details: We will use a Cox Proportional Hazards model with robust standard errors to estimate the hazard ratio and corresponding 95% confidence interval and Wald test statistic. Robust confidence intervals and p-values will be displayed. This analysis represents a subset selected post-randomization and therefore requires adjustment for confounders. Due to lack of known confounders early in the epidemic, we propose that the list of baseline descriptors be assessed individually for their association with the primary endpoint. Those associated univariably with $p<0.10$ will be included in this analysis. A hazard ratio greater than 1 (treatment/control) indicates efficacy of the intervention in shortening shedding.

Prevention of COVID-19 diagnosis

Objective: To test the efficacy of HCQ to prevent incident COVID-19.

Outcome: PCR-confirmed COVID-19 diagnosis post start of HCQ therapy through Day 14

Cohort(s): ITT, mITT, Sensitivity, Seronegative, Infected

Definition of Endpoint: COVID-19 is defined as symptomatic disease based on the following criteria with virologic confirmation (RNA PCR Ct $\leq$ 40):

- At least TWO of the following symptoms: Fever (≥ 38ºC), chills, rigors, myalgia, headache, sore throat, new olfactory and taste disorder(s), OR
- At least ONE of the following symptoms: cough, shortness of breath or difficulty breathing, OR
- Severe respiratory illness with at least 1 of the following:
    - Clinical or radiological evidence of pneumonia, OR
    - Acute respiratory distress syndrome (ARDS)

A participant will be defined as having reached the endpoint if they have the diagnosis above. The time of the endpoint will be the onset of symptoms. Participants who do not have the endpoint will be censored at the last time they were known to be symptom-free. Death and hospitalization will be counted in the endpoint.

Analysis Details: We will use a Cox Proportional Hazards model with robust standard errors to estimate the hazard ratio and corresponding 95% confidence interval and Wald test statistic. Efficacy will be calculated as 1 minus the hazard calculated from the treatment coefficient. Robust confidence intervals and p-values will be displayed. The following baseline variables will be included to increase precision: presence of metabolic disease (if present in at least 10% of the population), weight, sex, age, and type of contact.

7.1.3. Sensitivity efficacy analyses

The following sensitivity analysis will be performed including the one defined in Section 7.1.1.

Objective: To repeat the primary analysis redefining the endpoint to exclude inconclusive tests with Ct>38 from the definition of the endpoint.

Outcome: Polymerase chain reaction (PCR)-confirmed SARS-CoV-2 infection (self-collected samples collected daily Day 1 through Day 14])

Cohorts: mITT Cohort

Definition of Endpoint: A participant will be defined as having the endpoint if at least one of their PCR tests throughout follow-up is positive (both targets are identified or the Ct is <=38) or they develop confirmed COVID-19 through external evaluation (e.g. hospital or clinical assessment).

Analysis Details:

The same analytical approaches will be applied as for the primary efficacy analysis with a positive test now being defined as concordant results (positive) or a Ct <=38. Inconclusive tests will be treated as negative.

7.1.4. Stratified Analyses

All efficacy analyses will also be repeated within subgroups defined by contact type, site, symptomatic index case, household size, time to start intervention, quarantine status, duration of exposure to index case (8 hours or less vs more than 8 hours) and presence of children in household. These subgroup analyses will only be adjusted for the same baseline covariates as the primary analysis (excluding the variable used to define the subgroup). For each of these subgroups, the cumulative incidence of the primary endpoint of SARS-CoV-2 infection by 14 days will be estimated, including 95% confidence intervals within each strata and a 95% confidence interval of the difference in cumulative incidence at 14 days between strata.

8. INTERIM ANALYSIS AND DATA MONITORING COMMITTEE

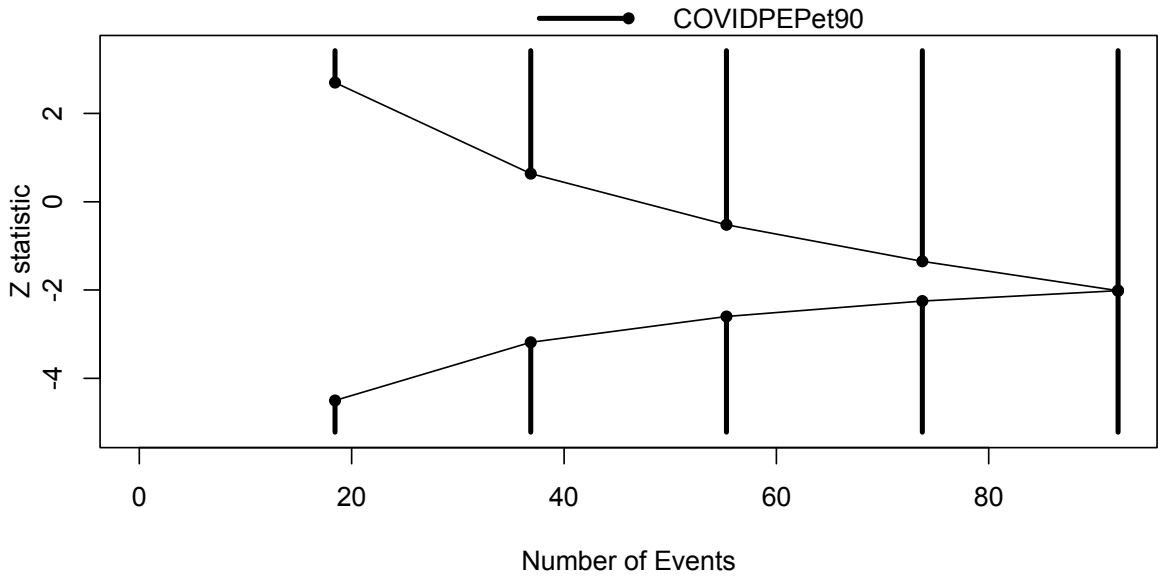
8.1. EFFICACY AND FUTILITY MONITORING

Interim monitoring will be performed on the primary outcome after approximately 20%, 40%, 60% and 80% of expected events. Monitoring will be performed using asymmetric O'Brien-Fleming bounds for early stopping for efficacy (lower) and futility (upper). The Wald test statistic from the robust Cox model without adjustment for covariates will be using for monitoring the trial with fraction of planned households as the information fraction[1].

Table 1 Monitoring bounds with 4 interim analyses. Expected events may vary and actual stopping bound will be based on number of events observed at the time of monitoring.

Analysis	Expected number of events	Efficacy bound	Futility bound (binding)
1	19	-4.50	2.7
2	37	-3.18	0.64
3	56	-2.60	-0.52

4	74	-2.25	-1.35
5	93	-2.01	-2.01



9. SAFETY ANALYSES

All safety analyses will be conducted on the ITT Cohort. The total number of serious adverse events (SAEs) and adverse events (AEs) will be summarized for each randomization arm by severity grade as shown in Mock Table 1. A line listing of SAEs will be presented by randomization arm. Attribution for SAEs and AEs as related or not related to the study medication as assessed by the clinical monitor will be reported. Social harms will be summarized for each randomization arm.

10. CHANGE HISTORY

This section will Identify major changes, if any after version 1.0 approval.

Version		Activity Description

Number	Effective Date	Affected Section(s)	

Mock Table 1. Total Number of Serious Adverse Events, by Severity Grade and Arm

	Randomized Arm					
	HCQ		Control		Total	
Severity Grade	**n**	**%**	**n**	**%**	**n**	**%**
Grade 1 / mild						
Grade 2 / moderate						
Grade 3 / severe						
Grade 4 / potentially life-threatening						
Grade 5 / death						

NOTE: The denominator for calculation of percentages is the number of SAEs at a given severity level with the exception of the 'Total' column, for which the denominator is the total number of SAEs.