Data Safety Monitoring Plan

with inclusion of Statistical Analysis Plan (SAP) in Section 6

for

WHIP COVID-19 Study

Revision Date

June 15, 2020

The second revision date

November 16, 2020 *with revised text as italic*

Table of Contents

1. Summary of the protocol

1.1 Overview

This is a prospective, multi-site study designed to evaluate whether the use of hydroxychloroquine in healthcare workers including healthcare providers and employees of hospitals or hospital systems (HCW), nursing home workers (NHW), employees of hospitals or hospital systems, first responders and correctional/law officers (FR), medical students (MS), public transit drivers, and family members of healthcare workers residing in the same dwelling, in the States of Michigan and Ohio can prevent the acquisition, symptoms and clinical COVID-19 infection.

The study will randomize a total of 3,000 participants who meet inclusion/exclusion criteria. The Study Participants will be randomized in a 1:1:1 blinded comparison of daily HCQ, weekly HCQ, or placebo. A fourth comparator group of HCW, NHW, MS, FR, public transit drivers, and family members of healthcare workers residing in the same dwelling who are already currently on standard HCQ therapy will be recruited to assess the impact of weight-based daily dosing of HCQ as compared to the randomized arms. Details are provided in the study protocol.

1.2 Primary and Secondary Aims

Primary Aim

1. To determine if the use of hydroxychloroquine as preventive therapy decreases the rate of acquisition of SARS-CoV 2 infections with clinical COVID-19 disease in Study Participants for each randomized treatment arm as compared to placebo.

Secondary Aims

1. Compare the rates of SARS-CoV 2 infections between the randomized treatment arms and the control arms to determine the effect of HCQ dose in the prevention of COVID-19 viremia and disease.

2. Compare the rates of SARS-CoV 2 infections in the non-randomized comparator arm to the randomized groups to assess the impact of chronic weight-based dosing of HCQ for COVID-19 prevention.

3. To compare the rate of SARS-CoV 2 infections as measured by SARS-CoV 2 serum

IgM/IgG in study participants receiving HCQ versus placebo.

4. To compare the seroprevalence of SARS-CoV 2 IgM/IgG positive samples at study entry and study conclusion in all participants receiving HCQ compared to those receiving placebo.

5. To compare the development of clinical symptoms or COVID-19 diagnosis in participants presenting asymptomatically at study entry but identified as seropositive by serology at entry between the randomized treatment arms and comparator arm.

6. To compare the clinical COVID-19 disease need for participants in each treatment arm to require emergency room visit, hospitalization, or ability to stay home without hospitalization.

7. To determine the safety and tolerability of HCQ dosing for preventive strategy against COVID-19.

8. To examine other clinical determinants contributing to the risk of SARS-CoV 2 infection including demographics, work type and location, positive COVID-19 partners, possible exposures and clinical symptoms.

9. Examine the association between HCQ drug levels and development of COVID-19 symptoms or positive test results.

10. To identify immunologic, serological and inflammatory markers associated with acquisition and response to COVID-19 in both HCQ and placebo Participants developing laboratory or clinical confirmed disease.

1.3 Inclusion/Exclusion Criteria

Inclusion Criteria

1. Participant is willing and able to provide informed consent.
2. Participant is 18-75 years of age.
3. Participant does not have symptoms of respiratory infection, including cough, fever (temperature >38.0 C), difficulty breathing, shortness of breath, chest pains, malaise, myalgia, headaches, nausea or

vomiting, or other symptoms associated with COVID-19.

4. Participant is willing and able to comply with study instructions, blood collection, other study procedures, and agrees to be contacted for required visits.
5. The participant has no known allergies or contraindications (as stated in the consent form) to the use of hydroxychloroquine (HCQ) as noted in the exclusion criteria and Pharmacy sections.
6. Participant is a HCW, FR, MS, NHW, public transit driver, or family member of a healthcare worker residing in the same dwelling, in the States of Michigan or Ohio at risk of exposure or exposed to COVID-19.

Exclusion Criteria

1. Does not meet inclusion criteria.
2. Participant has any of the symptoms attributable to or has currently tested positive for COVID-19 infection through RT-PCR nasal or oral-pharyngeal swab, a laboratory defined serology test, or an FDA approved assay.
3. Participant is currently taking select QT prolonging agents such as Lexapro, Celexa, or any prohibited medications per protocol. (See contraindicated medication list in Pharmacy Section)
4. Participant has a known co-morbidity including history of gastric bypass, epilepsy, cardiovascular disease, renal disease, or other organ failure that precludes the use of HCQ (see pharmacy section).
5. Participant has known renal failure with a creatinine clearance of <10 ml/min, pre-dialysis or requiring dialysis.
6. Participant is currently enrolled in another study to evaluate an investigational drug.
7. The participant is currently on diuretic therapy using loop-diuretics.
8. The participant was diagnosed with retinal disease or has history retinal surgery prior to study entry.
9. The participant has a diagnosis of porphyria prior to study entry.

10. The participant has a family history of Sudden Cardiac Death.
11. The participant has a history of known Prolonged QT Syndrome.
12. The participant has a history of ongoing liver failure.
13. A history of or ongoing serious illness or medical, physical, or psychiatric conditions(s), or any other factor that, in the Investigator's opinion, may interfer with the subject's ability to follow study procedures and complete the study.
14. Investigator site staff members directly involved in the conduct of the study.

Note that the inclusion/exclusion criteria may change during trial recruitment, after obtaining FDA and IRB approval.

1.4 Target Population, Sample Size and Study Duration

Approximately 3000 participants from the States of Michigan and Ohio who meet the study entry criteria will be prospectively enrolled and randomized at Henry Ford Health System (a member of the Detroit COVID Consortium). Future protocol amendments may expand the geographic areas from which patients are drawn and add additional enrolling sites. The fourth non-randomized comparator arm will include participants that are already on chronic HCQ maintenance therapy with weight-based, twice a day dosing for their autoimmune disease, which will be continued throughout the study. Enrollment in this group will be open as it will be a non-randomized treatment arm; the total number of participants enrolled in this group is expected to be lower than in the randomized groups. Total study duration is expected to be up to 2 years, with participant enrollment starting in April 2020.

1.5 Participating Enrolling Clinics & Data Coordinating Center

Clinical sites:

HFHS Center for Structural Heart Disease. This clinical site will be led by William W. O'Neill MD. and Dr. John E. McKinnon MD as the trial principle investigators (co-PIs), and coinvestigators (Co-Is) Dee Dee Wang, MD and Marcus Zervos, MD. The clinical site will be responsible for subject enrollment and follow-up. Other clinical sites may be added as needed for study recruitment.

Data Coordinating Center (DCC):

The HFHS Department of Public Health Science (PHS) will serve as the DCC. The DCC will provide statistical, database management, and site management expertise and support for the clinical trial—from study design, data collection and management, data analysis and reporting, to publication and dissemination. The DCC is also responsible for site staff training in data collection and management and will conduct site audits for data quality and study integrity and compliance with FDA regulations. Led by Dr. Mei Lu, director of the PHS DCC and a Biostatistician, the DCC staff include biostatisticians Jia Li PhD and Yueren Zhou MS, a project manager, Christina Melkonian MS, a project coordinator, Lora Rupp MS, and the database programmer, Yasmeen Hafeez.

A PHS operations committee, led by Dr. Christine C Johnson, Chair of HFHS PHS, and assisted by members Lonni Schultz PhD, biostatistician and head of Design and Biostatistics services, and Mr. Dan McLaren, Manager - Clinical Research Informatics Services of PHS, will oversee the trial conduct, including subject enrollment and data collection and management.

2. Study Data Collection & Quality Management

2.1 Case report forms (CRFs)

Case report forms (CRFs) that use Participants' study identifiers are designed and developed for the trial according to CDISC foundational standards as required for FDA-regulated trials. CRFs will be approved by the Trial PI. The DCC will work closely with the Trial PI and Biorepository Core to identify pre-analytic and analytic data elements to ensure data and specimen reproducibility and quality.

Due to the urgency for trial recruitment during the COVID pandemic period, paper CRFs were prepared and used in the beginning of study when the CRF database was not ready for the trial data entry. Trial recruitment began on April 10th using the paper CRFs.

2.2 Database management system for study data collection

The DCC has implemented a centralized, integrated electronic data capture (EDC) system—comprised of RedCAP, SQL, and Forte EDC data management systems—for trial data and sample collection and management: (1) RedCap is used to collect pre-screening data for possible eligible participants; (2) SQL is used for registering and tracking study samples; and (3) Forte EDC, an FDA 21 CFR Part 11-compliant data management system, is used for trial CRF data collection.

The DCC has prepared a data dictionary and all CRF fields will have consistent variable names to ensure accurate coding. A detailed Data Collection manual of procedures (MOP) has been prepared by the DCC as part of the training material for the clinical sites. The DCC will be responsible for training study staff in EDC as well as how to apply the MOP for CRF completion. The DCC will also provide a CRF data-quality and data-monitoring plan as part of the study documentation.

In general practice, a DCC CRF database should be ready prior to study enrollment. However, given the urgency for trial recruitment during the COVID pandemic period, paper CRFs were prepared and used before the Forte EDC database was built, validated and released on 4/17/2020. In addition, RedCap was used as an interim system to assist clinical sites for participant recruitment and study enrollment, and for the AE and SAE tracking log, until the Forte EDC database was up and running and clinical site staff were trained for CRF data entry. The RedCap AE and SAE tracking log was discontinued on 5/14/2020 when the Forte EDC AE and SAE tracking system was fully implemented. *To be able to enroll subjects elsewhere beside HFHS and to incorporate the study protocol revision to include pregnant women, a Forte EDC V2 database was implemented and went live on 7/26/2020 for new patient enrollment. The V1 database is referred to as the original database, which retained all patients enrolled prior to 7/26/2020 including their follow-up data collection. The V2 database has a revised CRF data cluster structure which is about 70% similar to the original Forte database (named V1 CRF data cluster, see Table in Appendix A). As a result, data quality checks are conducted separately for each version of the Forte database. The data will be integrated for analysis using SAS.*

2.3 EDC CRF data monitoring and quality management

A backlog of paper CRFs began to be entered into Forte EDC database on 4/20/2020 by PHS staff who had been trained for EDC data entry. That responsibility was transferred back to clinical study staff by the end of April.

Data can be entered directly into Forte EDC as they are collected at clinical sites to ensure a smooth and systematic flow of activities and to avoid backlogs and uneven workloads. Automated validation rules in Forte EDC will catch inconsistent data collection within or across CRFs as the data are entered. Forte EDC offers a user-friendly feature with color coding to reflect data status (data error versus error-free). A data discrepancy alert will also appear when a data element is improperly entered. For example, data element checks will be displayed as

soon as the data are entered into Forte EDC. Similar data validation processes can be implemented using SAS. For this study, both processes will be used.

A specimen ID will be issued once a specimen is collected and documented in both the CRF and specimen tracking databases. Specimen quality will be documented in the specimen tracking database after the specimen is registered in the SQL database at the DCC. Data will be seamlessly and automatically integrated daily to ensure timely data checks and data reports.

CRF Data quality reports will be generated daily and will include: (a) CRF status (missing but expected, incomplete but expected, complete as expected); and (b) data element status (missing, illogical, or invalid values; inconsistencies across CRFs in Forte EDC or using SAS).

The Clinical Site Coordinator will work closely with the DCC Data Coordinator to resolve data discrepancies by directly entering the missing-but-expected CRF or correcting data errors through Forte EDC. Data quality alerts will be generated continuously until resolved.

Prescreening data will be reported daily. The number of subject enrollments will be reported daily when data are available in the DCC central database.

Due to the volume of recruitment, a serious adverse event (SAE) report will be generated and reported to PIs and the DSMB weekly. The DCC has implemented an MOP for central SAE review and reporting process (detailed in a later section).

The DCC will work closely with the Trial PIs, laboratory core, and each site regarding non-compliant specimens, and action will be taken to remedy the issue if possible; the DCC will perform a descriptive analysis to compare specimen quality before and after the problem resolution.

3. Site Training and Management

Site staff training:

The DCC will assist the Trial PIs in the preparation of training materials and will conduct a series of role-specific training sessions for site personnel including the Trial PI, regulatory staff, and clinical coordinators on topics including the study protocol, Forte EDC training for patient scheduling, informed consent, and CRF data collection.

Training in Forte EDC will be performed at two levels: general Forte EDC data-entry training, and study-specific training. This training can be completed through webinars.

On- and Off-Site Visits:

DCC staff will discuss site visit results with the Trial PIs and clinical coordinators at the end of a site visit. An audit report will be written for each visit, listing all findings and requesting resolution. A follow-up meeting or conference call may be needed for clarification. Audit results will be documented by the DCC as part of the Trial Master Files (TMFs). The DCC will maintain close communication with the Trial PIs and Leadership Committee regarding auditing results. If a serious protocol violation occurs, action may be taken, as decided by the PI and DSMB.

4. Safety data collection and reporting

4.1 Adverse event definitions

For this study, the following standard AE definitions are used:

Adverse event: An Adverse Event (AE) is any untoward medical occurrence in a Participant administered an investigational product or undergoing a procedure, which does not necessarily have a causal relationship with the treatment. An AE can be any unfavorable and unintended sign, symptom, or disease temporally associated with the use of an investigational product, whether or not it is related to the investigational product.

Serious Adverse Event: A Serious Adverse Event (SAE) is any AE that results in any of the following outcomes:

- Results in death
- Is life-threatening
- Requires inpatient hospitalization or prolongation of existing hospitalization
- Results in persistent or significant disability/incapacity
- Is a congenital anomaly/birth defect.

AEs are graded according to the following scale:

Mild: An experience that is transient and requires no special treatment or intervention. The experience does not generally interfere with usual daily activities. This includes transient laboratory test alterations.

Moderate: An experience that is alleviated with simple therapeutic treatments. The experience impacts usual daily activities. Includes laboratory test alterations indicating injury, but without long-term risk.

Severe: An experience that requires therapeutic intervention. The experience interrupts usual daily activities. If hospitalization (or prolongation of hospitalization) is required for treatment it becomes an SAE.

The study uses the following AE attribution scale:

Not related: The AE is clearly not related to the study drug (i.e., another cause of the event is most plausible, and/or a clinically plausible temporal sequence is inconsistent with the onset of the event).
Possibly related: An event that follows a reasonable temporal sequence from the initiation of study procedures, but that could readily have been produced by a number of other factors.
Related: The AE is clearly related to the study drug.

Each SAE will be reviewed by an independent medical monitor, Dr. Bojana Jovanovic MD, a rheumatology specialist at HFHS, who is not be involved in the proposed study. The detailed review process is detailed in MOP.

4.2 Safety reporting

All problems having to do with subject safety will be reported by the Principal Investigator (PI) to the IRB within ten working days. Specifically, the following will be reported, in writing: 1) all serious adverse events potentially associated with the study drug, and/or 2) any incidents or problems involving the conduct of the study or patient participation, including problems with the recruitment and/or consent processes. The PI will provide a discussion of any side effects or problems noticed in the course of the study to the IRB.

The DCC will produce safety reports that list serious adverse events, deaths, and disease or treatment specific events, in aggregate, for DSMB weekly meetings or at a scheduled meeting.

The DSMB will review safety reports by blinded treatment group. If there are a significant number of adverse events, the DSMB may request that the treatment groups be unblinded to ensure there are not problematic side effects.

5. Reports for Study Progress and Quality Control

5.1 Study progress Reports

The DCC will produce administrative reports that describe study progress. These reports will be reviewed internally by the DCC and PI for ongoing quality control and will be presented to the Data Safety Monitoring Board (DSMB) at scheduled meetings.

5.2 Subject accrual / randomization Reports

The DCC will produce daily accrual / randomization reports, which will include:

1) The number of prospective Participants who completed a Volunteer Survey, and their eligibility, demographics and occupations.

2) The number of Participants currently enrolled.

A weekly report to DSMB will include a cumulative report of enrollment and enrollment projection.

5.3 Data collection and management Reports

Data quality will be assessed using measures listed below:

- CRF status by form (overall and by clinical site): the number of CRFs received the number of expected CRFs, and percentage of missing expected CRFs.
- The number of data queries generated and the number/percent of queries resolved (overall and by clinical site).
- Protocol deviations including: 1) number of subjects who were ineligible but randomized into the trial; 2) numbers of subjects who were enrolled but not according to randomization sequence.

5.4 DSMB Reports

A weekly report to DSMB will include 1) Cumulative report of enrollment and projection; 2) AE and SAE report.

Reports for a scheduled DSMB meeting will include the cumulative daily and weekly reports; safety report; and the efficacy results at the interim analysis session.

6. Statistical Analysis Plan (SAP)

A formal SAP will be developed before study completion or database lock for an interim analysis. Statistical issues not addressed in the study protocol will be addressed in the formal SAP.

6.1 Study Design

Participants who meet study entry criteria and are not on HCQ prior to study enrollment will be randomized in a 1:1:1 blinded comparison of daily or weekly hydroxychloroquine versus placebo for 8 weeks. The randomization will be stratified by enrolling site and high- vs. low- risk category, where the high-risk category will consist of HCWs who work in COVID-19 care areas.

A fourth non-randomized comparator group comprised of study participants who are already currently on standard HCQ therapy for an autoimmune disease(s) will be used to assess the impact of weight-based daily dosing of HCQ as compared to the randomized arms.

The trial will use a pragmatic, adaptive enrichment design, in which the primary efficacy analysis will be based on all randomized patients with a negative screen for COVID-19 at baseline, whereas the safety analysis will be based on all randomized patients, including those who are later determined to have been positive for COVID-19 at baseline (given that there may be a time lag to get initial screening COVID-19 test results back after study enrollment).

The trial is designed to have one interim look— when 50% of subjects have completed the study and primary endpoint assessment—for an early treatment efficacy assessment, and possible reassessment of study sample-size if the estimation of placebo infection rate is inadequate.

6.2 Locking Data for Analysis/Transfer

Data will be locked (frozen) for each planned data transfer/ formal data analysis. After the last patient's data/ specimens are collected and all data discrepancies are resolved or closed, data can be locked. The final data analysis/ transfer will be completed within 30 days after the last patient's data/ specimen collection. The data transfer process includes documentation of the database structure, data dictionary with details regarding data element names, labels, formats, and distribution, as well as documentation regarding data collection details, as needed.

6.3 Study Aims, Hypothesis and Proposed Analysis

*6.3.1 **Primary** Aim and Hypothesis*

As noted below, the primary aim has been changed to: (based on Study Protocol dated 05/26/2020)

To determine if the use of hydroxychloroquine as preventive therapy decreases the rate of acquisition of SARS-CoV 2 infections with clinical COVID-19 disease in Study Participants for each randomized treatment arm as compared to placebo.

From: (based on original Study Protocol, dated at and after 04/06/2020)

To determine if the use of hydroxychloroquine as preventive therapy decreases the rate of acquisition of SARS-CoV 2 infections in Study Participants as compared to placebo.

Primary Hypothesis (P1): Hydroxychloroquine administered daily decreases the rate of acquisition of SARS-CoV 2 infections with clinical COVID-19 disease, compared to placebo; and hydroxychloroquine administered weekly decreases the rate of acquisition of SARS-CoV 2 infections with clinical COVID-19 disease, compared to placebo.

The **primary outcome** will be defined as clinical COVID-19 disease with laboratory confirmation of infection.

Presentation with COVID-19 symptoms during study follow-up (defined as fever ≥38° C, or two or more non-fever symptoms that are new since the baseline) will refer participants to obtain SARS-CoV 2 testing through employee health or participants' primary care physician, with test results documented in their medical record.

Laboratory confirmation of SARS-CoV 2 infection (SARS-CoV 2 test positive) will be defined as at least one positive RT PCR, or both IgM and IgG positive serology. The lab test results can be obtained from (a) study blood samples (IgM and IgG serology), and (b) RT PCR test results ordered by a participant's primary care physician or through employee health.

All participants who test SARS-CoV 2 positive will be followed for at least 30 days after the first positive test for presentation of symptoms.

6.3.2 Full Analysis Set (FAS)

All randomized subjects who do not have a SARS-CoV 2 positive test at the baseline, have taken at least one dose of study medication and provide at least one post-baseline follow-up during the 8-week treatment period. Participants will be analyzed according to the treatment group to which they were assigned according to Intention to treat (ITT).

6.3.3 Safety Analysis Set:

The safety analysis set will consist of all subjects randomized who receive at least one dose of study medication.

6.3.4. Proposed Analysis

6.3.4.1 General Statistical Consideration:

The statistical software package SAS® Version 9.4 or higher (SAS Institute, Cary, NC, USA) will be used to perform all analyses and to summarize data.

The analyses described in this statistical analysis plan (SAP) are considered a priori, in that they have been defined prior to database lock. Any analyses not described herein performed subsequent to database lock will be considered post hoc and exploratory.

6.3.4.2. Handling of missing data:

For summary statistics, results will be reported based upon observed data. The study will make all possible efforts to obtain patient's follow-up data and laboratory confirmation of infection for COVID 19 disease determination. The efficacy analysis will be performed based on all observed data assuming missing as random (MAR), two selected imputation approaches are specified for the sensitivity analysis.

6.3.4.3 Primary Efficacy Analysis.

The FAS is considered. Chi-square test will be used to compare the rate of COVID-19 disease in each randomized treatment group to that of the placebo group using Dunnett's step-up method to adjust for multiple comparisons and controlling for type I error[1,2]. Specifically, Dunnett's step-up method uses a closed gatekeeping strategy, in which the largest *p* value will be evaluated first at critical value of 0.05; if that fails, the second *p* value will be evaluated at criterial value of 0.0262 (details in the sample size /power calculation section). A treatment efficacy is detected if the largest *p* value <0.05 or the second *p* value<0.0262. The results of

efficacy analysis will be validated through two specified imputation approaches described in below

Sensitivity Analysis #1: A Tipping Point Analysis (TPA) with delta-adjusting imputation

The primary analyses for COVID 19 disease is based on the observed data and the results are valid if the missing data are missing at random (MAR), meaning that the probability of a value being missing, conditional on the observed data and factors in the statistical model, is random and not dependent on the unknown value of the missing data point.

To address this, we will conduct Tipping Point Analysis (TPA) under the condition of missing not at random (MNAR) to assess the robustness of the primary analysis. Specifically, the TPA first step (STEP 1) is to conduct m (e.g., m=5 or 10) multiple imputation (MI) based on baseline covariates, and the prior to 8-week and post-baseline covariates to estimate the missing primary binary endpoint (D) with the underlying assumption of MAR. We will use the Markov Chain Monte Carlo (MCMC) algorithm based on the assumption of multivariate normality (Schafer 1997, Allison 2001) for estimating the mean of D (Allison 2017) stratified by treatment group without rounding to binary values. However, this approach is still under the MAR assumption and hence the result will be further examined by shifting the imputed mean of D in a range of negative to positive values, which can be done by inflecting or suppressing the mean of D among those who had missing D in treated group. A tipping point is identified if it reverses the study conclusion. Those results will be further evaluated on whether the imputed value associated with the tipping point is plausible. SAS Procedures PROC MI and PROC MIANALYZE and SAS %midata macro can be used for TAP analysis.

Sensitivity Analysis #2: A control-based MI

For the most conservative approach, MI can be restricted using data only from the placebo group for the missing endpoint D in both the treated and placebo groups using a similar approach described above for the sensitivity analysis #1 with the shift=0 (the first step).

6.3.4.4 Secondary Aims, Hypotheses and Analyses

S1: To compare the rates of SARS-CoV 2 infection status between the randomized treatment arms and the control arms to determine the effect of HCQ dose in the prevention of COVID-19 viremia and disease.

Secondary Hypothesis (S1): There is a hydroxychloroquine dose effect on acquisition of SARS-CoV 2 infection status, compared to placebo.

The endpoint will be SARS-CoV 2 infection status, defined as 1) no SARS-CoV 2 test positive during study follow up; 2) SARS-CoV 2 test positive (i.e., at least one positive RT PCR, or both IgM and IgG positive serology) without clinical symptoms ("asymptomatic SARS-CoV 2 test positive") 3) SARS-CoV 2 test positive with clinical symptoms ("symptomatic SARS-CoV 2 test positive").

The analysis will be conducted using FAS. Ordinal logistic regression will be used and analysis will start testing the overall daily and weekly dose effect, followed by pairwise/subgroup comparisons if an overall dose effect is detected at 0.05 level.

S2: To compare the rates of SARS-CoV 2 infection status in the non-randomized comparator arm to the randomized groups to assess the impact of chronic weight-based dosing of HCQ for COVID-19 prevention.

Secondary Hypothesis (S2): There is weekly /daily HCQ dose effect on acquisition of SARS-CoV 2 infection, compared to patients who were already on HCQ.

The endpoint will be the same as described for S1. The analysis will include all randomized *in FAS and non-randomized study participants with at least one post-baseline follow-up.*

We will first test for treatment selection bias using the data collected at the baseline. Inverse probability treatment weighting (IPTW)/propensity score, or matched sub-cohort will be considered to adjust for imbalanced participant characteristics. A similar analysis to that of S1 will be used for testing hypothesis S2.

S3: To compare the rate of SARS-CoV 2 infections as measured by serum IgM/IgG in participants receiving HCQ versus placebo.

Secondary Hypothesis (S3): There is weekly /daily dose effect on status of SARS-CoV 2 infection as measured by serum IgM/IgG, compared to placebo.

The endpoint will be serum IgG/IgM status at each participant's study conclusion, classified as:

1. Negative (consistently negative for both serologies)
2. Indeterminate (IgM positive, with IgG consistently negative)
3. Positive (IgM and IgG positive, or IgG positive)

The analysis will be conducted *using FAS*. Ordinal logistic regression will be used to study the status distribution differences among three Arms. Similar to the analysis for S1, the analysis will start testing the overall dose effect, followed by the pairwise comparison between each treated arm and placebo.

S4: To compare the seroprevalence of SARS-CoV 2 IgM/IgG positive samples at study entry and study conclusion in all participants receiving HCQ compared to those receiving placebo.

Secondary Hypothesis (S4): There is weekly /daily dose effect on change in SARS-CoV 2 seroprevalence at study conclusion from the baseline, compared to placebo.

Serum IgM/IgG status at baseline and at study conclusion will be classified into the three categories described for S3. **The endpoint** will be a change in serum IgM/IgG status (from negative at baseline to indeterminate or positive at study conclusion, or indeterminate at baseline to positive at study conclusion).

The analysis approach described for S1 will be used to study the dose effect on change in serum IgM/IgG status *based on all randomized subjects*.

S5: To compare the development of clinical symptoms or COVID-19 diagnosis in participants presenting asymptomatically at study entry but identified as seropositive by serology at entry between the randomized treatment arms and comparator arm.

Secondary Hypothesis (S5): There is dose effect on development of clinical symptoms or COVID-19 diagnosis.

The endpoint will be the same as that defined for P1. This analysis will include all randomized study participants who, at baseline, have positive serology (IgM or IgG) but are clinically asymptomatic. The analysis approach described for S1 will be used for this analysis.

S6: To compare the clinical COVID-19 disease need for participants in each treatment arm to require emergency room visit, hospitalization, or ability to stay home without hospitalization.

Hypothesis (S6): There is dose effect on incidence of ED visit/hospitalization, compared to placebo.

The endpoint will be rate of COVID-related ED visits or hospitalizations (by end of the planned 8-week follow-up period, or within 30 days after SARS-CoV 2 test positive if the positive test occurs <30 days before the end of the planned 8-week follow-up period).

FAS will be considered. The analysis approach described for S1 will be used for this analysis.

S7: To determine the safety and tolerability of HCQ dosing for preventive strategy against COVID-19.

Secondary Hypothesis (S7): There is an HCQ dose effect on AEs and SAEs, compared to placebo.

The endpoint will be any AE and any SAE observed during the 8 weeks of follow-up after enrollment.

Safety Analysis Set is considered by including a*ll randomized subjects.*

A descriptive analysis will report AE frequency and grade distribution among the three arms, and include a detailed list of serious adverse events (SAEs).

S8: To examine other clinical determinants contributing to the risk of SARS-CoV 2 infection including demographics, work type and location, positive COVID-19 partners, possible exposures and clinical symptoms.

Secondary Hypothesis (S8): Risk for SARS-CoV 2 infection varies by subject demographics, work type and location, and clinical conditions.

The possible covariates will include subject demographics, work type and location, and clinical conditions at the baseline prior to study enrollment, as well as the study treatment status. The endpoint will be clinical COVID-19 disease confirmed by laboratory test (same endpoint as P1).

FAS is considered. Logit univariate and multivariate modeling will be conducted with consideration of possible treatment-by-baseline covariate interaction. Variables will be retained in the final model if there is significant individual effect or variable-by-treatment interaction at p-value 0.05. Area under the receiver operating characteristic curve (AUROC) will be used to assess the predictive ability.

S9: To examine the association between HCQ drug levels and development of COVID-19 symptoms or positive test results.

Secondary Hypothesis (S9): There is an association between HCQ drug levels and development of COVID-19 symptoms or positive test results.

The endpoint will be the same, as described for S1 *using FAS*.

A similar analysis as that described for S1 will be used. The analysis will include concentration of HCQ in blood at baseline, week 4 and week 8 as covariates.

S10: To identify immunologic, serological and inflammatory markers associated with acquisition and response to COVID-19 in both HCQ and placebo Participants developing laboratory or clinical confirmed disease.

Secondary Hypothesis (S10): There are immunologic, serological and inflammatory markers for COVID-19 positive.

Two separate endpoints, SARS-CoV 2 test positive, and COVID19 disease with SARS-CoV 2 test positive laboratory confirmation (P1), will be considered.

FAS is considered. The immunologic, serological and inflammatory markers will be collected using the baseline blood samples. The same analysis approach as described for S8 will be used for S10. The baseline variables identified from S8 will be included for S10 analysis.

6.4 Sample size/power calculation

Power Calculation for original Primary Aim: (based on original Study Protocol, dated 04/06/2020)

To determine if the use of hydroxychloroquine as preventive therapy decreases the rate of acquisition of SARS-CoV 2 infections in Study Participants as compared to placebo.

The sample size and power calculation for the original aim are based on a SARS-CoV 2 infection rate of 10% in the placebo-treated group, and an average of 6.8% in the randomized weekly- and daily- HCQ groups, with a relative reduction of 32%. Based on an alpha=0.05 two-sided test, with 900 subjects per group that are COVID negative at the baseline, and one interim analysis when 50% of participants (450 per group) have completed the study and primary

endpoint data assessment, we will have more than 80% power (82% to be exact using nQuery) to detect a significant HCQ effect if the observed test statistic is ±2.9626 or p-value<0.003 in the interim analysis, or ±1.9685 or p-value<0.0497 in the final analysis, using an O'Brien-Fleming alpha spending method to ensure an overall type 1 error of 0.05 and assuming that overall COVID positive rate in the weekly- and daily- dose HCQ groups is 6.8% with n=1800 in the HCQ treated group and 900 in the placebo treated group. A reduction of 36% can be detected between either one of the HCQ treatment groups and placebo, using the same assumptions. Assuming 5-10% of patients will have SARS-CoV 2 positive at the baseline, we need 1000 per group, with a total of 3000 participants to complete the trial.

Power Calculation for revised Primary Aim: (based on revised Study Protocol dated 04/16/2020 or after)

> To determine if the use of hydroxychloroquine as preventive therapy decreases the rate of acquisition of SARS-CoV 2 infections with clinical COVID-19 disease in Study Participants for **each** randomized treatment arm as compared to placebo.

The revised sample size and power calculation are based on a 10% rate of clinical COVID-19 disease (with laboratory confirmation) in the placebo-treated group, and an average 6.8% rate in the two randomized HCQ-treated groups (weekly and daily dosing groups), with a relative reduction of 32%. We propose two looks; one look (interim analysis) when 50% of the participants have completed the study and primary endpoint assessment (for P1), and a final look when 100% of participants have completed the study.

Based on an alpha=0.05 two-sided test, chi-square test/logistic regression with odds ratio estimates, and Haybittle-Peto alpha spending method to ensure an overall type I error, the critical value for the interim analysis will be 0.001 and the critical value for the final analysis will be 0.05. For data illustration, we have also included an O'Brien-Fleming spending function approach for controlling type I error with a critical value of 0.0054 for the interim analysis and 0.0492 for the final analysis.

Power calculations are based on n=900 or 935 participants per study arm, using a simulation approach (n=10,000 simulations) in three correction methods: #1 comparing the weekly and daily HCQ dose arms to the placebo group separately using the Dunnett's step-up approach; #2 comparing the weekly and daily HCQ dose arms to placebo separately using a conservative Bonferroni approach with critical value of 0.025 for each test at the final look; #3 comparing the combined weekly and daily HCQ dose arms to the placebo using a single test.

With Haybittle-Peto boundary and alpha=0.05 at the final look, the Dunnett's step-up correction uses a closed gatekeeping strategy for testing the daily dose arm vs placebo and the weekly dose arm vs placebo separately: the largest *p* value tested first will be at 0.05; if that fails, the second test will be at *p* value 0.026, delivered from the last row of Table 2 of Dunnett and Tamhane (1992), and validated based on simulation using R *DunnettTests* package. Likewise, at the interim look, the largest *p* value tested first will be at 0.001; if that fails, the second test will be at 0.0005.

Table 1. Power for different scenarios using varying methods. Sample size was determined using the Dunnett's Step up method with the Haybittle-Peto boundary.

Haybittle-Peto approach	p_0=0.1, p_1=0.073, p_2=0.063		p_0=0.1, p_1=p_2=0.068	
Method	n=900	n=935	n=900	n=935
#1 Dunnett's Step up	0.800	0.811	0.770	0.800
#2 Bonferroni	0.784	0.800	0.745	0.760
#3 Daily+weekly arms vs placebo	0.812	0.83	0.819	0.835

p_0 is the rate for the placebo treated group and p_1 and p_2 are the rates for the two treated groups.

O'Brien-Fleming approach	p_0=0.1, p_1=0.073, p_2=0.063		p_0=0.1, p_1=p_2=0.068		
Method	n=910	n=935	n=900	n=935	n=970
#1 Dunnett's **Step up**	0.8	0.814	0.76	0.776	0.801
#2 Bonferroni	0.787	0.8	0.741	0.761	0.784
#3 Daily+weekly arms vs placebo	0.815	0.83	0.815	0.83	0.846

p_0 is the rate for the placebo treated group and p_1 and p_2 are the rates for the two treated groups.

For the primary Aim / hypothesis, we will use Dunnett's Step up approach to test each dose effect vs. placebo, and a sample size of 935 per arm will insure at least 80% power. The power will be higher for some of the secondary Aims (S1-S4 if use the compressed binary endpoint of infection presence or absence), since a single test will be used to study overall dose effect first, followed by subgroup analysis (Correction Method #3)

Assuming 5-10% of patients will be excluded from FAS, we need 1000 subjects per group, with a total of 3000 participants to complete the trial. The power and sample size calculation remained unchanged.

6.4.1 Sample size reassessment

The adaptive design will allow the study to reassess the rate of *SARS-CoV 2 infection with clinical COVID-19 disease* observed in the placebo-treated group at the interim look.

6.5 Interim analysis schedule

One interim analysis will be conducted when 50% of subjects have completed the study and primary endpoint assessment.

6.6 Preparing the data for analyses

The interim analysis database will be locked 1 week after the cutoff date. The data are received continually throughout the trial and checked for consistency. Queries will be issued on an ongoing basis. The project statistician will forward statistical programs and data to the DSMB statistician before the scheduled DSMB meeting. The DSMB statistician will replace the dummy random treatment code with the actual treatment code and execute the programs. The DSMB statistician will summarize the findings in a report addressed to the other members of the DSMB.

6.7 Stopping rule

- For the interim analyses of the primary efficacy analysis, Haybittle-Peto boundary stopping rule will be used. Treatment efficacy will be determined, compared to placebo, if the largest p-value is less than 0.001, or if that fails, the second p-value is less than 0.0005.
- DSMB will consider stopping the trial if there is a safety concern to study participants.

7. Data Safety Monitoring Board

Board Membership: The DSMB will be appointed locally for this investigator-initiated study. Members will include:

DSMB Chair/Director: Sid Goldstein, MD HFHS Cardiovascular Medicine
Tel: 313-916-2720 Email: SGOLDST1@hfhs.org

Jerry Yee, MD HFHS Division Head-Nephrology
Tel: (313)916-9405 Email: jyee1@hfhs.org

Erica Herc, MD HFHS Infectious Disease
Tel:(313)-850-8694 Email: eherc1@hfhs.org

Amita Bishnoi, MD, HFHS Department of Internal Medicine Rheumatology
Tel: 313-433-6244 (cell phone) Email: ABISHNO1@hfhs.org

George Divine, PhD HFHS Department of Public Health Sciences (Biostatistician)
Tel: 313-799-1494 (cell phone) Email: gdivine1@hfhs.org

The members of the DSMB must be experts in the field, not be involved in the study, and have no conflicts of interest with the study/research.

Safety oversight will be under the purview of the DSMB. The DSMB will meet to assess safety on each arm of the study, to protect the interests of research patients and ensure that they are not exposed to undue risk. The DSMB will evaluate the progress of the trial with regard to data quality, recruitment, accrual and retention, and risk vs. benefit to participants.

The DSMB chair will be provided with a weekly site AE and SAE report to monitor safety. An interim analysis will be done when a half of patients have completed the 8 weeks assessments for an early treatment efficacy or a possible reassessment of study sample-size if the estimation of placebo rate is inadequate. Depending on response to the study intervention, the DSMB may decide it is necessary to stop/continue the trial if 1) there are safety concerns at any point in the trial or 2) there is evidence of treatment efficacy at the interim analysis; or recommendation for trial modification as it may be suggested, or 3) to continue the trial as designed. The DSMB will only stop the clinical trial for safety reasons.

8. Confidentiality

8.1 Protection of Subject Privacy

During this study, a limited physical examination will be performed, and questionnaires will be administered. Data will be kept in strict confidence. No information will be given to anyone outside of the study without permission from the subject. This statement guarantees confidentiality. Confidentiality is assured by use of identification codes. All data, whether generated in the laboratory or at the bedside, will be identified with a randomly generated identification code unique to the subject in accordance with Health Information Portability and Accountability Act (HIPAA) guidelines.

8.2 Database Protection

The database is secured with password protection. The informatics manager receives only coded information, which is entered into the database under those identification codes. Electronic communication with outside collaborators will involve only unidentifiable information.

8.3 Confidentiality during AE Reporting

AE reports and annual summaries will not include subject-identifiable material. Each will include the identification code only.

Appendix A: Forte V1 and V2 CRF Data Clusters

Databases	Data Cluster V1 (4/17/2020)	Data Cluster V2 (7/21/2020)
	informcons	informcons
	EligibilityForm	Eligibility
Eligibility	OperatEligibility	~~OperatEligibility~~
	Demographics	Demographics
	OccupSocialHist	OccupSocialHist
	BriefMedHist	BriefMedHist
Baseline	UpperRespiratorySymptoms	
Randomization	RandomNonRandom	Randomization Health_Questionnaire
	BodyMeasure	BodyMeasure new
	BloodSplRst	BloodSplCollection
Blood Sample Results	OperatBloodSplRst	BloodSplRst
	UpperRespiratorySymptoms	Health_Questionnaire
	StatUpdateQuest	StatUpdateQuest New
Wk 1-8	OperatStatQuest	~~OperatStatQuest~~
Concomitant meds	ConcomitMed	ConcomitMed
AE Log	AdverseEvent	AdverseEvent
Subject Drug Accountability Log	Drug Accountability	Drug Accountability
Protocol Deviation Log	Deviation	Deviation new
FinalStatus	Final Status	Final Status
Unblinding	Treatment Unblinding	Treatment Unblinding
Call Log	Telephone contact	Telephone log
COVID Status (7/9/20202)	**COVID Status**	**COVID Status**
Final COVID_Status (7/9/2020)	**Final COVID_Status**	**Final COVID_Status**
Pregnancy		Pregnancy_Form, Pre or Post Partum CRFs
Study Notes		Notes
Fetal/Neonatal AE		Fetal/Neonatal_AE

Forte Database

- 2 new CRFS were added on 7/9/2020 to capture COVID event data
- 75% of data the CRF data clusters in V2 Forte database are new

Redcap Database

- E-consent
- Self-scheduling for baselin survey and blood drawn
- Drug shipment
- Randomization
- Clockwise for follow-up

The system is in production **(live on 7/26/2020)**

Red--new/modified CRF data clusters

References

1. Allison, Paul D. (2001) Missing Data. Thousand Oaks, CA: Sage Publications.
2. Horton N.J., Lipsitz S.R., and Parzen, M. (2003) A potential for bias when rounding in multiple imputation. American Statistician 57: 229-232.
3. Schafer, Joseph L. (1997) Analysis of Incomplete Multivariate Data. London: Chapman & Hall.
4. Paul D. Allison, University of Pennsylvania (2017) Multiple Imputation of Categorical Variables Under the Multivariate Normal Model
5. Charles W. Dunnett and Ajit C. Tamhane. A step-up multiple test procedure. Journal of the American Statistical Association, 87(417):162-170, 1992.
6. Howard DR, Brown JM, Todd S, Gregory WM. Recommendations on multiple testing adjustment in multi-arm trials with a shared control group. Stat Methods Med Res. 2018 May;27(5):1513-1530. doi: 10.1177/0962280216664759. Epub 2016 Sep 19.