PROTECT Trial
STATISTICAL ANALYSIS PLAN

May 1, 2020

Version 1.0

TABLE OF CONTENTS

1 ABBREVIATIONS

Abbreviation	Description
AZM	Azithromycin
DSMB	Data Safety Monitoring Board
HCQ	Hydroxychloroquine
SOC	Standard of Care

2 INTRODUCTION

The purpose of this document is to describe statistical analysis plans for the PROTECT Trial interim and final analyses. It includes an overview of the project, description of endpoints, general analysis conventions, analysis populations, handling of missing data, statistical analyses, and guidelines for interpreting interim analysis results. The statistical team intends to perform all statistical analyses as specified in the protocol. However, it is conceivable that unanticipated changes to the course of the study may preclude some scheduled analyses. Study conditions or analysis results may identify a need for additional data review.

3 STUDY DESIGN AND OBJECTIVES

PROTECT is an open-label, factorial randomized clinical trial with two factors and nested randomization. Patients hospitalized with confirmed SARS-CoV-2 infection will be initially randomized to one of two arms in a 1:1 ratio: 1) Standard of care treatment or 2) Standard of care treatment plus 5 days of hydroxychloroquine (HCQ). Patients with no contraindication to azithromycin (AZM) who have not received AZM in the seven days prior to enrollment will undergo a second randomization in a 1:1 ratio to either receive 1) No AZM or 2) 5 days of AZM. The resulting treatment group combinations are summarized in Table 1. Participants will be followed for 28 days after randomization or until hospital discharge, whichever is later. The primary endpoint is the World Health Organization (WHO) ordinal scale for clinical improvement as measured on day 14 after randomization (see Section 6.1 for details).

Table 1. Treatment groups

	AZM: No	**AZM: Yes**	**AZM: Not Eligible**
HCQ: No	Group 1: SOC	Group 3: SOC + AZM	Group 5: SOC, AZM-ineligible
HCQ: Yes	Group 2: SOC + HCQ	Group 4: SOC + HQC + AZM	Group 6: SOC + HCQ, AZM-ineligible

HCQ = hydroxychloroquine, AZM = azithromycin, SOC = standard of care
Hypothesis #1 (HCQ versus no HCQ): Groups 2+4+6 versus 1+3+5
Hypothesis #2 (AZM versus no AZM): Groups 3+4 versus 1+2

3.1 Primary Objective

The PROTECT trial aims to simultaneously test the following two primary hypotheses:

- **Hypothesis #1 (HCQ versus no HCQ):** The addition of HCQ improves outcomes as measured by the day 14 WHO ordinal scale (primary endpoint) when added to standard of care and possible concomitant AZM. The analysis will be a comparison of groups 2+4+6 versus 1+3+5 as defined in Table 1.

- **Hypothesis #2 (AZM versus no AZM):** The addition of AZM improves outcomes as measured by the day 14 WHO ordinal scale (primary endpoint) when added to standard of care and possible concomitant HCQ. The analysis will be a comparison of groups 3+4 versus 1+2 as defined in Table 1.

3.2 Secondary Objectives

The trial will also evaluate the HCQ and AZM hypotheses using the following secondary endpoints:

- Mortality during the index hospitalization
- Number of days on mechanical ventilation during the index hospitalization
- Progression to mechanical ventilation during index hospitalization
- World Health Organization ordinal scale on day 28
- Days from randomization until hospital discharge
- Duration of fever after randomization
- Duration of supplemental oxygen use after randomization

3.3 Tertiary Objectives (Exploratory)

The trial will perform additional exploratory pairwise treatment group comparisons for the primary and secondary endpoints as follows (see Table 1 for definitions of study groups):

- SOC + HCQ versus SOC (group 2 versus 1)
- SOC + AZM versus SOC (group 3 versus 1)
- SOC + HCQ + AZM versus SOC (group 4 versus 1)
- SOC + AZM versus SOC + HCQ (group 3 versus 2)
- SOC + HCQ + AZM versus SOC + HCQ (group 4 versus 2)
- SOC + HCQ + AZM versus SOC + AZM (group 4 versus 3)

4 RANDOMIZATION METHODOLOGY

Patients meeting entry criteria will be stratified by eligibility to undergo randomizations for both study drugs or only HCQ.

<u>Stratum 1: Eligible for both study drugs</u>
Participants who are eligible for both study drugs will be randomized in a 1:1:1:1 ratio to one of the following:

Group	Description

1	Standard of Care
2	Standard of Care + HCQ
3	Standard of Care + AZM
4	Standard of Care + HCQ + AZM

Stratum 2: Ineligible for azithromycin

Participants who are ineligible to be randomized to azithromycin will be randomized in a 1:1 ratio to one of the following groups

Group	Description
5	Standard of Care --
6	Standard of Care + HCQ

Randomizations within each stratum will use permuted blocks with random block sizes and will be further stratified by ventilator status at baseline.

5 POPULATIONS FOR EFFICACY AND SAFETY ANALYSES

5.1 Intention-to-treat population

The intention-to-treat (ITT) population will include all participants who receive a randomized treatment assignment, including those who are later determined not to meet trial eligibility criteria, unless the randomization was an unintentional administrative error, as determined by the coordinating center at the Duke Office of Clinical Research. Patients will be assigned to treatment groups based on each patient's randomized treatment group assignment regardless of which treatment if any the patient actually received.

6 DESCRIPTION OF EFFICACY ENDPOINTS

6.1 Primary Endpoint

The primary outcome measure is the World Health Organization (WHO) ordinal scale for clinical improvement as measured on day 14 after randomization. The WHO scale has 9 levels numbered 0 to 8 with 0 denoting no clinical or viroligic evidence of infection and 8 denoting death.

Table 2. World Health Organization ordinal scale for clinical improvement

Descriptor	Score
No clinical or virologic evidence of infection	0
No limitation of activities	1
Limitation of activities	2
Hospitalized, no oxygen therapy	3
Oxygen by mask or nasal cannula	4
Non-invasive ventilation or high-flow oxygen	5
Intubation and mechanical ventilation	6
Mechanical ventilation + additional organ support (pressors, renal replacement therapy, ECMO)	7
Death	8

7 GENERAL ANALYSIS CONSIDERATIONS

Analyses will be performed according to the principle of "intention-to-treat" with subjects analyzed (and endpoints attributed) according to the group to which enrolled subjects were randomized, regardless of subsequent medications or treatment crossover. All major treatment group comparisons will be performed using Bayesian statistical methods. The output of a Bayesian analysis is a posterior probability distribution describing the relative likelihood of different numerical estimates for unknown quantities. This posterior distribution can be used to determine the likelihood of a clinically important treatment benefit or harm in light of the study data. The Bayesian approach provides an especially useful perspective for interpreting results of small studies that are likely to be accompanied by substantial statistical uncertainty. The Bayesian analog of a confidence interval, known as a credible interval, has a direct and intuitively appealing interpretation as an interval containing the true value with a specified probability. Point estimates will be calculated as posterior means and accompanied with 95% equal-tailed credible intervals. Analysis results will also be framed in terms of probability statements, for example, the probability that the study drugs lead to improved outcomes on the primary endpoint. The analysis plan does not include frequentist hypothesis testing and does not make adjustment for multiple comparisons. Additional considerations related to multiple comparisons are discussed in Sections 11.1.

7.1 Missing Data

Every effort will be made to obtain complete data. In the case where missing data remain, the following procedures will be followed. These procedures are based on the assumption that missing data will be rare and may be revised if that assumption is violated.

- **Primary endpoint.** If a day 14 WHO scale assessment is unavailable then analysis will be based on the closest assessment within 21 days of randomization with ties broken by taking the earlier assessment. If a patient is known to have died or been re-hospitalized but the exact date is unknown then data entry may be based on an estimated value. If no estimate of a patient's status is available within 21 days of randomization then the primary endpoint will be regarded as missing and will not be imputed. The approach to handling missing primary endpoint data may be revised if >5% of patients have no assessments within a window of 14 ± 3 days or >2% of patients have no assessments within a window of 14 ± 6 days.
- **Mortality during the index hospitalization.** Missing in-hospital mortality status will not be imputed.
- **Number of days on mechanical ventilation during index hospitalization.** No imputation will be performed for this endpoint.
- **Progression to mechanical ventilation during index hospitalization.** No imputation will be performed for this endpoint.
- **Day 28 WHO scale.** If unavailable then analysis will use the closest measurement from days 14 to 40 with ties broken by taking the later assessment. Data entry may be based on estimated dates if exact dates for a status assessment is unavailable.
- **Days from randomization until hospital discharge.** No imputation will be performed for this endpoint.
- **Duration of fever after randomization.** No imputation will be performed for this endpoint.
- **Duration of supplemental oxygen use after randomization.** No imputation will be performed for this endpoint.

7.2 Pooling Across Trial Sites

All data from all sites will be pooled for analysis. Study site will be ignored in the primary analysis of the primary endpoint but will be included as a covariate in a planned covariate-adjusted secondary analysis.

7.3 Withdrawals, Dropouts, Loss to Follow-up

All patients and all data will be included in the analyses. Patients who withdraw from the study or are lost to follow-up before collection of the primary endpoint will be analyzed using their last available follow-up data according to the imputation rules described in Section 7.1. Participants with no analyzable follow-up data for any safety or efficacy endpoints will be excluded and will not appear in tabulations of baseline data. The number of such exclusions, if any, will be reported.

8 ANALYSIS OF THE PRIMARY ENDPOINT

8.1 Parametric Win Ratio Analysis

The primary endpoint analysis will focus on assessing the overall effects of standard of care plus HCQ versus no HCQ (Hypothesis #1) and AZM versus no AZM (Hypothesis #2), as detailed in Section 3.1. Analysis will be performed in the intention-to-treat (ITT) population, as defined in Section 3.1. The analytic approach will be a model-based parametric implementation of the win ratio framework. [1] In a conventional win ratio analysis, the win ratio statistic is calculated by forming all possible pairs of 1 participant from the treatment group and 1 participant from the control group and then dividing the proportion of pairs for which the treated patient has a better outcome by the proportion of pairs for which the untreated patient has a better outcome. The PROTECT trial analysis will estimate this same underlying win ratio quantity but will do so parametrically using a statistical model. The parametric approach was chosen because it facilitates analysis in a Bayesian statistical framework using a range of Bayesian prior distributions. A Bayesian nonparametric implementation of the win ratio with a non-informative prior distribution was tested in simulations prior to trial launch and was found to be anti-conservative in the sense that the prior distribution assigned high prior probability to the hypothesis of large or small win ratio and this led to high probability of a spurious signal during interim analyses. The underlying quantities to be estimated in this analysis are as follows:

Win Ratio for Hypothesis #1 (HCQ versus no HCQ): The win ratio for HCQ versus no HCQ is defined as the probability that a randomly selected patient who was randomized to HCQ has a better outcome than a randomly selected patient from the same AZM subgroup who was randomized to no HCQ divided by the probability that a randomly selected patient who was randomized to HCQ has a worse outcome than a randomly selected patient from the same AZM subgroup who was randomized to no HCQ.

Win Ratio for Hypothesis #2 (AZM versus no AZM): The win ratio for AZM versus no AZM is defined as the probability that a randomly selected patient who was randomized to AZM has a better outcome than a randomly selected patient from the same HCQ subgroup who was randomized to no AZM divided by the probability that a randomly selected patient who was randomized to AZM has a worse outcome than a randomly selected patient from the same HCQ subgroup who was randomized to no AZM.

8.2 Calculation Formulas for the Parametric Win Ratio

The WHO primary endpoint is a categorical variable with 9 categories labeled 0 to 8 where 0 represents no clinical or virological evidence of infection and category 8 represents death. Because there are 9 possible values, the probability distribution of the primary endpoint within each treatment group can be described by a set of 9 probabilities which sum to 1. We will use notation π_{gj} to denote the probability that the primary endpoint of a patient in group g takes the value j (see Table 3). Using this notation, the probability that a patient in group g has a better outcome (win) than a patient in group h can be calculated by the expression:

$$W_{g:h} = W_{g:h}(\boldsymbol{\pi}) = \Pr\left(Y^g < Y^h\right) = \sum_{j=0}^{7} \Pr(Y^g = j) \Pr\left(Y^h > j\right) = \sum_{j=0}^{7} \pi_{gj}\left(\pi_{h,j+1} + \cdots + \pi_{h,8}\right)$$

and this expression holds for all $g, h = 1,2, \ldots ,6$. We use the notation $W_{g:h} = W_{g:h}(\boldsymbol{\pi})$ to emphasize that $W_{g:h}$ is implicitly a function of the unknown π_{gj}'s as defined in Table 3. The win ratio comparing groups g versus h is then given by

$$\text{Win-Ratio}_{g:h} = \frac{W_{g:h}}{W_{h:g}}.$$

Hypothesis #1 (HCQ). The win ratio for Hypothesis #1 is a stratified version of the win ratio. It first randomly selects a patient from one of the HCQ groups (groups 2, 4, or 6) and then randomly selects a comparison patient from the corresponding non-HCQ group within the same AZM stratum, such that group 2 is paired with group 1, group 4 is paired with 3, and group 6 is paired with 5. The win ratio is the probability that the patient receiving HCQ has a better outcome than the patient not receiving HCQ divided by the probability that the patient receiving HCQ has a worse outcome than the patient not receiving HCQ. Mathematically, this turns out to be equivalent to a weighted average of 3 pairwise win ratios for groups 2 versus 1, 4 versus 3, and 6 versus 5. Specifically, the calculation formula for the hypothesis #1 win ratio is

$$\text{Win-Ratio}^{\text{HQC}} = \left(\frac{n_1 + n_2}{n}\right)\frac{W_{2:1}}{W_{1:2}} + \left(\frac{n_3 + n_4}{n}\right)\frac{W_{4:3}}{W_{3:4}} + \left(\frac{n_5 + n_6}{n}\right)\frac{W_{6:5}}{W_{5:6}}$$

where $n_g$is the number of participants in the g-th treatment stratum and $n = n_1 + \cdots + n_6$ is the total sample size.

Hypothesis #2 (AZM). The win ratio for Hypothesis #2 is a stratified version of the win ratio. It first randomly selects a patient from one of the randomized AZM groups (groups 3 or 4) and then randomly selects a comparison patient from the corresponding non-AZM group within the same level of HCQ, such that group 3 is paired with 1 and group 4 is paired with 2. The win ratio is the probability that the patient receiving AZM has a better outcome than the patient not receiving AZM divided by the probability that the patient receiving AZM has a worse outcome than the patient not receiving AZM. Mathematically, this turns out to be equivalent to a weighted average of 2 pairwise win ratios for groups 3 versus 1 and 4 versus 2. Specifically, the calculation formula for the hypothesis #1 win ratio is

$$\text{Win-Ratio}^{\text{AZM}} = \left(\frac{n_1 + n_2}{n_1 + n_2 + n_3 + n_4}\right)\frac{W_{2:1}}{W_{1:2}} + \left(\frac{n_3 + n_4}{n_1 + n_2 + n_3 + n_4}\right)\frac{W_{4:3}}{W_{3:4}}.$$

Table 3. Notation for defining the distribution of WHO categories within each treatment group

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
0	π_{10}	π_{20}	π_{30}	π_{40}	π_{50}	π_{60}
1	π_{11}	π_{21}	π_{31}	π_{41}	π_{51}	π_{61}

2	π_{12}	π_{22}	π_{32}	π_{42}	π_{52}	π_{62}
3	π_{13}	π_{23}	π_{33}	π_{43}	π_{53}	π_{63}
4	π_{14}	π_{24}	π_{34}	π_{44}	π_{54}	π_{64}
5	π_{15}	π_{25}	π_{35}	π_{45}	π_{55}	π_{65}
6	π_{16}	π_{26}	π_{36}	π_{46}	π_{56}	π_{66}
7	π_{17}	π_{27}	π_{37}	π_{47}	π_{57}	π_{67}
8	π_{18}	π_{28}	π_{38}	π_{48}	π_{58}	π_{68}

Note: $\pi_{g1} + \pi_{g2} + \cdots + \pi_{g8} = 1$ for all $g = 1,2, \ldots,6$.

A point estimate of the win ratio can be calculated using the "plug-in" method by substituting (i.e. plugging in) data-based estimates $\hat{\pi}_{gj}$ in place of the unknown π_{gj}'s in the calculation formulas given above. This leads to estimates of the form

$$\widehat{\text{Win-Ratio}}^{\text{HQC}} = \left(\frac{n_1 + n_2}{n}\right)\frac{\widehat{W}_{2:1}}{\widehat{W}_{1:2}} + \left(\frac{n_3 + n_4}{n}\right)\frac{\widehat{W}_{4:3}}{\widehat{W}_{3:4}} + \left(\frac{n_5 + n_6}{n}\right)\frac{\widehat{W}_{6:5}}{\widehat{W}_{5:6}}$$

$$\widehat{\text{Win-Ratio}}^{\text{AZM}} = \left(\frac{n_1 + n_2}{n_1 + n_2 + n_3 + n_4}\right)\frac{\widehat{W}_{2:1}}{\widehat{W}_{1:2}} + \left(\frac{n_3 + n_4}{n_1 + n_2 + n_3 + n_4}\right)\frac{\widehat{W}_{4:3}}{\widehat{W}_{3:4}}$$

$$\text{where} \quad \widehat{W}_{g:h} = \sum_{j=0}^{7} \hat{\pi}_{gj}\left(\hat{\pi}_{h,j+1} + \cdots + \hat{\pi}_{h,8}\right), \quad g, h = 1,2,3,4,5,6,$$

and $\hat{\pi}_{gj}$ is a data-based estimator of π_{gj}. For example, in a simple (non-Bayesian) nonparametric analysis we could plug in $\hat{\pi}_{gj} = n_{gj}/n_g$.where n_{gj} is the number of participants in group g with a WHO score equal to j. The primary analysis method is similar to this but uses Markov Chain Monte Carlo methods for inference and will estimate π_{gj}'s parametrically using a statistical model as defined in Section 8.3.

8.3 Statistical Model

The trial's primary endpoint is a numeric variable taking the values 0 to 8 where 0 represents no clinical or virological evidence of infection and category 8 represents death. To account for this variable's discreteness and reduce dependence on modeling assumptions, the outcome will be analyzed as an ordered categorical variable without assuming a specific parametric distribution. The primary analysis will be performed using a parametric cumulative logits (aka proportional odds) model. The proportional odds methodology is an extension of binary logistic regression to accommodate $K > 2$ ordered outcome categories. The model considers all possible ways of dichotomizing K categories into a binary outcome and assumes that the set of all possible binary outcomes is related to covariates through a set of ordinary binary logistic regression models with shared regression coefficients. The specific form of the model is

$$\text{logit}\,\Pr(Y \le k|H, A, x) = \begin{cases} \theta_k & \text{if } H = 0, A = 0 \quad (\text{group } 1) \\ \theta_k + \mu & \text{if } H = 1, A = 0 \quad (\text{group } 2) \\ \theta_k \quad + \gamma_1 & \text{if } H = 0, A = 1 \quad (\text{group } 3) \\ \theta_k + \mu + \gamma_1 + \delta_1 & \text{if } H = 1, A = 1 \quad (\text{group } 4) \\ \theta_k \quad + \gamma_2 & \text{if } H = 0, A = 2 \quad (\text{group } 5) \\ \theta_k + \mu + \gamma_2 + \delta_2 & \text{if } H = 1, A = 2 \quad (\text{group } 6) \end{cases}, \quad k = 0,1, \ldots,7$$

where $\text{logit}\, p = \log p/(1-p)$, Y denotes the outcome, H is an indicator of HCQ randomization assignment (0 = randomized to no HCQ, 1 = randomized to HCQ), A is an indicator of AZM randomization assignment (0 = randomized to no AZM, 1 = randomized to AZM, 2 = N/A, not eligible for AZM

randomization), the θ_k's ($k = 0,1, \ldots ,7$) are intercepts describing the cumulative frequency of WHO categories within each treatment group, and μ, γ_1, γ_2, δ_1,δ_2 are main effects and interaction parameters describing outcome differences across treatment groups. The terms δ_1 and δ_2 are interaction terms and they describe the extent to which the treatment effect of randomization to HCQ versus no HCQ to differs depending on AZM group assignment and vice versa. Sample size calculations for PROTECT were based on fitting a model that omits interaction effects and assumes that $\delta_1 = \delta_2 = 0$. Under the assumed model we have the following expressions for the probabilities defined in Table 1:

Table 4. Expressions for cell probabilities as a function of proportional odds model parameters

$$\pi_{1j} = \frac{\exp(\theta_j)}{1+\exp(\theta_j)} - \frac{\exp(\theta_{j-1})}{1+\exp(\theta_{j-1})}$$

$$\pi_{2j} = \frac{\exp(\theta_j + \mu)}{1+\exp(\theta_j + \mu)} - \frac{\exp(\theta_{j-1} + \mu)}{1+\exp(\theta_{j-1} + \mu)}$$

$$\pi_{3j} = \frac{\exp(\theta_j + \gamma_1)}{1+\exp(\theta_j + \gamma_1)} - \frac{\exp(\theta_{j-1} + \gamma_1)}{1+\exp(\theta_{j-1} + \gamma_1)}$$

$$\pi_{4j} = \frac{\exp(\theta_j + \mu + \gamma_2 + \delta_2)}{1+\exp(\theta_j + \mu + \gamma_2 + \delta_2)} - \frac{\exp(\theta_{j-1} + \mu + \gamma_2 + \delta_2)}{1+\exp(\theta_{j-1} + \mu + \gamma_2 + \delta_2)}$$

$$\pi_{4j} = \frac{\exp(\theta_j + \mu + \gamma_1 + \delta_1)}{1+\exp(\theta_j + \mu + \gamma_1 + \delta_1)} - \frac{\exp(\theta_{j-1} + \mu + \gamma_1 + \delta_1)}{1+\exp(\theta_{j-1} + \mu + \gamma_1 + \delta_1)}$$

$$\pi_{5j} = \frac{\exp(\theta_j + \gamma_2)}{1+\exp(\theta_j + \gamma_2)} - \frac{\exp(\theta_{j-1} + \gamma_2)}{1+\exp(\theta_{j-1} + \gamma_2)}$$

$$\pi_{6j} = \frac{\exp(\theta_j + \mu + \gamma_2 + \delta_2)}{1+\exp(\theta_j + \mu + \gamma_2 + \delta_2)} - \frac{\exp(\theta_{j-1} + \mu + \gamma_2 + \delta_2)}{1+\exp(\theta_{j-1} + \mu + \gamma_2 + \delta_2)}$$

for $j = 0,1, \ldots 8$, where $\theta_{-1} = -\infty$ and $\theta_8 = \infty$.

8.4 Estimation Methods

Model parameters will be estimated in a Bayesian statistical framework using Markov Chain Monte Carlo sampling for inference [2]. Advantages of the Bayesian framework include the ability to express analysis results in terms of clinically relevant probabilities and the ability to perform exact inference on complicated functions of model parameters, for example, the win ratio.

Prior distribution

Bayesian analysis requires the specification of a probability distribution representing prior information about the set of unknown model parameters before observing the study data. Because no single choice of prior can accurately capture the viewpoints of all potential stakeholders, the analysis will be performed with multiple choices for the prior distribution. For the primary analysis, the prior for treatment main and interaction effects will be a set of independent normal distributions:

$$\mu \sim N(\text{mean} = 0, \text{SD} = 0.3537) \quad \text{(main effect)}$$
$$\gamma_1 \sim N(\text{mean} = 0, \text{SD} = 0.3537) \quad \text{(main effect)}$$
$$\gamma_2 \sim N(\text{mean} = 0, \text{SD} = 0.3537) \quad \text{(main effect)}$$
$$\delta_1 \sim N(\text{mean} = 0, \text{SD} = 0.2651) \quad \text{(interaction)}$$
$$\delta_2 \sim N(\text{mean} = 0, \text{SD} = 0.2651) \quad \text{(interaction)}$$

The main effects SD = 0.3537 was chosen to imply that odds ratios for HCQ versus no HCQ and AZM versus no AZM are between 0.5 and 2.0 with 95% probability assuming that the interaction effects are zero. The interaction SD = 0.2651 was chosen to imply that the ratio of odds ratios for HCQ versus no HCQ across two levels of AZM and the ratio of odds ratios for AZM versus no AZM across two levels of HCQ would each be between 0.8 and 1.25 with 80% probability. The intercept parameters $\theta_k$will be assigned central t distributions with 3 degrees of freedom and SD = 10 subject to the constraint $\theta_0 < \theta_2 < \cdots < \theta_7$, which is the default prior in the R package 'brms'. To explore sensitivity to the choice of prior, the analysis will be repeated using two alternative forms for the prior distribution. The first alternative prior distribution will be the same as above but will set $\delta_1 = \delta_2 = 0$ with 100% probability. The second alternative prior will assign a non-informative improper uniform distribution to the set of main effects and interaction terms. Additional prior distributions may be implemented post-hoc for manuscript presentations or by request of external reviewers and stakeholders.

Computation

Posterior means and credible intervals for the Win Ratio quantities of interest will be calculated using Hamiltonian Markov Chain Monte Carlo (MCMC) simulations as implemented in the R package 'brms'. [4,4] To reduce Monte Carlo error, we will generate a large number of simulated parameter values and will perform graphical and other diagnostic checks to ensure convergence. To perform inference on $\text{Win-Ratio}^{\text{HQC}}$ and $\text{Win-Ratio}^{\text{AZM}}$ we note that each of these quantities is a function of the cell probabilities $\boldsymbol{\pi}$ defined in Table 3 which in turn are functions of the model parameters $\Theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_7, \theta_6, \theta_7, \mu, \gamma_1, \gamma_2, \delta_1, \delta_2)$. We use notation $\text{Win-Ratio}^{\text{HQC}}(\Theta)$ and $\text{Win-Ratio}^{\text{AZM}}(\Theta)$ to denote the value of the win ratios as a function of unknown model parameters. Posterior mean point estimates of the win ratios will be calculated as

$$\widehat{\text{Win-Ratio}}^{\text{HCQ}} = \frac{1}{M}\sum_{i=1}^{M} \text{Win-Ratio}^{\text{HCQ}}(\Theta^{(i)})$$
$$\widehat{\text{Win-Ratio}}^{\text{AZM}} = \frac{1}{M}\sum_{i=1}^{M} \text{Win-Ratio}^{\text{AZM}}(\Theta^{(i)})$$

where M is the number of MCMC iterations and $\Theta^{(i)}$ is the simulated value of Θ on the i-th iteration of the MCMC procedure. A 95% credible for the win ratios will be obtained by calculating the 2.5th and 97.5th empirical percentiles of the quantities $\text{Win-Ratio}^{\text{HCQ}}(\Theta^{(i)}$ and $\text{Win-Ratio}^{\text{AZM}}(\Theta^{(i)})$ across the M simulated values. Finally, the probability that each win ratio falls above or below a clinically relevant threshold will be calculated by counting the proportion of simulated win ratio values that fall above or below the threshold.

8.5 Secondary Analyses of the Primary Endpoint

The primary analysis is unadjusted for baseline covariates but a covariate-adjusted analysis will also be presented for the major hypotheses in order to target a potentially more clinically relevant treatment effect estimand. Baseline covariates for this analysis will include ventilator status at the time of

randomization and enrolling site. To implement the covariate-adjusted analysis, baseline ventilator status and site will be included as main effects covariate adjustments in the proportion odds model along with the set of treatment group indicators used in the primary analysis. Win ratios for HCQ and AZM will be defined in a manner analogous to the primary analysis. The covariate-adjusted win ratio for HCQ will be a comparison between a hypothetical randomly selected patient from one of the HCQ groups (2, 4, or 6) and a hypothetical randomly selected patient from the same site, same baseline ventilator status, and same AZM subgroup who was randomized to no HCQ. The covariate-adjusted win ratio for AZM will be a comparison between a hypothetical randomly selected patient from one of the AZM groups (3 or 4) and a hypothetical randomly selected patient from the same site, same baseline ventilator status, and same HCQ subgroup who was randomized to no AZM.

8.6 Contingency for Violation of Proportional Odds Assumption

The proportionality assumption underlying the proportional odds model will be assessed by examining plots of empirical cumulative logits across treatment groups and comparing observed versus model-predicted probabilities. If there is substantial evidence of an unacceptable degree of model misspecification (e.g. the empirical cumulative logits from different treatment groups are clearly crossing) then a nonparametric instead of parametric win ratio analysis will be presented. To implement a Bayesian nonparametric version of the win ratio analysis, cell probabilities underlying the win ratio calculation (Table 3) will be estimated empirically by assigning noninformative independent Dirichlet distributions to the set of category probabilities within each treatment group. Other details of the calculations are essentially identical to those described for the primary analysis. The prior for this nonparametric Bayesian analysis was chosen for simplicity and computational convenience and may be revised in a subsequent version of the Statistical Analysis Plan.

8.7 Exploratory Pairwise Group Comparisons

The treatment effects estimated by the primary analysis are averages in the sense that the comparison of HCQ versus no HCQ is aggregated over AZM subgroups and the comparison of AZM versus no AZM is aggregated over HCQ subgroups. The primary reason for targeting average as opposed to subgroup-specific treatment effects which are potentially more informative for practice is the expected lack of statistical precision for estimating subgroup-specific effects. In order to shed light on potential treatment effect heterogeneity across AZM and HCQ subgroups, we will report estimated odds ratios and accompanying 95% credible intervals comparing treatment groups on a pairwise basis from the primary analysis proportional odds model.

9 ANALYSIS OF SECONDARY EFFICACY ENDPOINTS

The major treatment group comparisons of HCQ versus no HCQ and AZM versus no AZM will be repeated using the following secondary clinical endpoints in place of the primary endpoint:

- Mortality during the index hospitalization
- Number of days on mechanical ventilation during the index hospitalization
- Progression to mechanical ventilation during index hospitalization.
- Day 28 WHO ordinal scale
- Days from randomization until hospital discharge
- Duration of fever after randomization
- Duration of supplemental oxygen use after randomization

Analysis of **index hospitalization mortality status** and **mechanical ventilation** will be performed using logistic regression. The analysis will use the same predictors as the primary endpoint proportional odds model and will use the same Bayesian prior distribution for the main effects and interaction effects parameters. For the ventilation endpoint, patients on ventilator at baseline will be excluded. To facilitate interpretation, parameter estimates from the logistic regression model will be re-expressed as win ratios, risk ratios, and/or risk differences.

Analysis of the **day 28 WHO scale** will be identical to the primary endpoint analysis using the day 28 assessment in place of the 14 day assessment.

Analysis of **length of hospital stay, duration of fever, and duration of supplemental oxygen use** will be based on a continuation ratio model for discrete time-to-event data treating death as a competing risk. For length of stay, the continuation ratio approach models that probability of being discharged on day $d+1$ conditional on being alive and in-hospital on day d and assumes that covariates have a multiplicative effect on this probability. For patients who die prior to discharge, the patient's discharge date will be set to the last date of the patient's potential follow-up time. The interpretation is similar for duration of fever and supplemental oxygen. Predictor variables in the continuation ratio model will be identical to the primary endpoint proportional odds model and will use a similar prior distribution for the main effects and interaction terms.

10 PRESENTATION OF SAFETY DATA

Safety data will be collected for all randomized participants from the time of randomization through day 28 unless the participant withdraws from the study before day 28. The information to be collected on safety-related events includes event narratives, investigator-assigned likelihood of relatedness to study treatments (unlikely, possibly, probably), distinctions between serious and non-serious adverse events, expectedness of the event, and the event outcome. Event severity will be classified using the Common Terminology Criteria for Adverse Events (CTCAE) v3.0. Tabulations of these data will be performed in the intention to treat population using the treatment groups defined in Section 5.1. Safety data will be tabulated by treatment group and will include the following.

- Frequency and percentage of serious adverse events (SAEs) overall, by treatment group, and by various cross classifications of treatment group, CTCAE classification, expectedness, relatedness to study treatments, and outcome.
- Frequency and percentage of all-cause study drug discontinuation overall and by treatment group
- Frequency and percentage of each of the following adverse events of special interest (EOSI) overall and by treatment group:
    - Arrhythmias (ventricular), not including torsade de pointes
    - Torsade de pointes
    - Hepatic Failure
    - Bone marrow failure
    - Aplastic anemia
    - Prolonged QT interval
    - Angioedema
    - Exfoliative dermatitis
    - Acute generalized exanthematous pustulosis (AGEP)

    - Psychosis
    - Suicidal Ideation
    - Seizure
- Narrative descriptions of unanticipated problems
- Patient- and event-level listings and additional tabulations as requested by the DSMB

11 INTERIM ANALYSES

The primary objective of interim analyses is to ensure the safety of the participants enrolled in the trial and to evaluate the accumulating safety and efficacy data by treatment group to test for possible differences favoring any of the randomized treatments or treatment combinations. In addition, interim monitoring will involve a review of participant recruitment, compliance with the study protocol, status of data collection, an assessment of whether control group event rates are consistent with the rates hypothesized in the sample size calculations, and other factors which reflect the overall progress and integrity of the study including potential geographic differences.

An independent DSMB will review safety data on a regular basis according to the schedule specified in the DSMC charter. A team of one or two independent statisticians (i.e., the DSMB reporting statisticians) will be the only individuals (aside from the DSMB) with knowledge of interim outcomes by treatment group. Safety listings and tabular summaries will be generated regularly for the DSMB review. Unplanned meetings may be called if the DSMB members deem it to be necessary. The frequency of generating/sending these tables and listings to the DSMB is specified in the DSMC Charter.

Interim efficacy analyses will focus on the primary endpoint (World Health Organization ordinal scale) and death during the index hospitalization. Analyses will be performed both for the comparison of hydroxychloroquine versus no hydroxychloroquine (Hypothesis #1) and for the comparison of azithromycin versus no azithromycin (Hypothesis #2). The first pre-planned interim review of efficacy data will be performed after completion of data collection for the primary endpoint of the first 75 participants and subsequent pre-planned efficacy reviews will be performed approximately bi-weekly or based on a schedule specified in the DSMB charter. Considerations for multiple testing of the accumulating trial data are discussed in Section 11.1.

Statistical methods for interim analyses will be identical to those described in Section 8 for the final analysis. For the WHO scale primary endpoint, the statistical framework will be a Bayesian parametric win ratio analysis based on the proportional odds model. For mortality, the statistical framework will be a Bayesian logistic regression analysis. Guidelines proposed to the DSMB for interpretation of interim results are summarized in Table 5:

Table 5. Proposed guidelines for DSMB statistical monitoring

	Definition	**Monitoring Guideline**
Moderate efficacy.	Win Ratio ≥1.25 favoring study drug over no study drug	Signal if probability is >80%
Any efficacy	Win Ratio >1.0 favoring study drug over no study drug	Signal if probability is >95%

Inefficacy	Win Ratio <1.0 favoring no study drug over study drug	Signal if probability is >80%
Moderate harm	Win Ratio <0.9 favoring no study drug over study drug	Signal if probability is >75%

Note: Study drug refers to the comparison of HCQ versus no HCQ (hypothesis #1) or the comparison of AZM versus no AZM (hypothesis #2). Win ratios will be estimated for both the primary endpoint and death during the index hospitalization.

11.1 Multiplicity considerations for interim analyses

The repeated analysis of multiple endpoints across multiple time points gives rise to the possibility of drawing an incorrect conclusion as a result of the play of chance. Based on Monte Carlo simulations performed during protocol development, we estimate that the probability of a signal occurring for moderate or any efficacy if outcomes are monitored continuously across 500 patients and there is truly no between group difference in the primary endpoint is 14%. The Bayesian approach to controlling the probability of an incorrect conclusion involves selecting a prior distribution that faithfully reflects the user's prior degree of skepticism concerning the possibility of a clinically meaningful treatment benefit or harm. If the Bayesian prior distribution is well calibrated to the user's degree of skepticism, then the calculated posterior probabilities are appropriate to use for decision making without the requirement for a multiple comparisons adjustment. Although we have specified a set of guidelines for monitoring by the DSMB, we do not specify a degree of Bayesian certainty that should be regarded as "definitive evidence" at the end of the trial. Instead, we assume that stakeholders will use the calculated probabilities to inform their decision making and that different users may apply different thresholds for determining whether the degree of certainty is sufficient to warrant any particular conclusion or action based on the data.

12 REFERENCES

1. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. European heart journal. 2012 Jan 1;33(2):176-82.
2. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis. CRC press; 2013 Nov 1.
3. Bürkner P (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1–28. doi: 10.18637/jss.v080.i01
4. Stan Development Team. 2018. RStan: the R interface to Stan. R package version 2.17.3. http://mc-stan.org

STATISICAL ANALYSIS PLAN REVISION HISTORY

Version	Date	Description of Changes
1.0	May 1, 2020	Original SAP