Azithromycin for COVID-19 Treatment In Outpatients Nationwide (ACTION)

Statistical Analysis Plan

1. Administrative Information

Trial registration:	ClinicalTrials.gov NCT04332107
Funder:	Bill & Melinda Gates Foundation (grant # INV-017026)
SAP Version:	Version 6, Updated 2020-5-29
	A revision history for this document is included at the end.
Protocol Version:	Refers to Protocol Version 1.29
Contributors:	Statisticians responsible:
	Benjamin F. Arnold [1] , Travis C. Porco [1]
	Principal Investigators:
	Catherine E. Oldenburg [1], Thuy Doan[1]

[1] Francis I. Proctor Foundation for Research in Ophthalmology, University of California, San Francisco, USA

SAP Table of Contents

The content of this Statistical Analysis Plan meets the requirements stated by the US Food and Drug Administration and conforms to the American Statistical Association's Ethical Guidelines. This SAP was organized following guidelines proposed in:

Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318: 2337–2343. PMID: 29260229 [1]

The companion computational notebook with underlying sample size calculations presented herein is entitled: `ACTION-sample-size-power-symptom-free.Rmd / .html`; it is saved in the same directory as this document.

2. Introduction

2.1. Background and rationale

Identification of a safe, effective treatment for individuals with mild or moderate COVID-19 that prevents disease progression and reduces hospitalization would reduce the burden on the health system. High dose hydroxychloroquine is being evaluated for SARS-CoV-2 prevention and COVID-19 disease treatment, but has a high risk of a number of potentially severe adverse events. Recent evidence has indicated that the broad-spectrum macrolide azithromycin may have some activity against coronaviruses. A large community randomized trial in Niger demonstrated reduced viral load among children with commensal coronaviruses (prior to the emergence of SARS-CoV-2) in communities receiving biannual mass azithromycin distribution compared to placebo (submitted). In SARS-CoV-2 patients in France, the addition of azithromycin to a hydroxychloroquine regimen appeared to decrease SARS-CoV-2 positivity by PCR compared to hydroxychloroquine-only and control patients. Azithromycin is generally well-tolerated and may be an attractive option for treating patients with mild disease.

2.2. Objectives

1: To determine if a single oral dose of azithromycin is effective for reducing progression to hospitalization in patients with a positive SARS-CoV-2 PCR test with mild or moderate disease.

3. Study Methods

3.1. Trial design

The trial will be a scalable, telemedicine-based parallel group simple trial to evaluate the efficacy of azithromycin for the prevention of COVID-19 disease progression and hospitalization.

3.2. Randomization

Participants will be randomized 2:1 to either receive a dose of oral azithromycin or oral placebo. Patients will be twice as likely to receive azithromycin than placebo. The randomization will not be stratified or blocked in any way. We found that even with permuted block sizes and ≥6 labels (4 azithromycin, 2 placebo), a 2:1 allocation could lead to unmasking through longer runs of active treatment labels.

3.3. Sample size

The trial's revised target enrollment is 455 participants, which will be sufficient to detect an increase in the proportion of patients who are symptom free 50% to 65% (15 percentage point increase) at 14 days from enrollment.

With 455 participants, the trial will have 80% power to detect an increase of 15 percentage points in the percentage of patients who are symptom free at 14-d under the following assumptions:

- 2:1 allocation ratio of azithromycin : placebo
- 50% symptom-free under placebo at the 14-day measurement (binomial outcome)
- 20% loss to follow-up
- 2-sided alpha of 0.05, power = 0.8.

- A single interim analysis with 50% of primary outcomes collected with stopping guideline determined using the Lan-DeMets spending function approach with an O'Brien and Fleming boundary.

The table below shows sample sizes under several effect size scenarios, assuming azithromycin increases the proportion of patients who are symptom free by 0.1 to 0.2 compared to placebo.

Sample size required for a fixed design allowing for 20% loss to follow-up. Calculations assume 2-sided alpha of 0.05 and 80% power.

p placebo	p azithro	RD	N placebo	N azithro	N Total
0.5	0.60	0.10	348	695	1,043
0.5	0.61	0.11	287	573	860
0.5	0.62	0.12	240	479	719
0.5	0.63	0.13	204	407	611
0.5	0.64	0.14	175	349	524
0.5	0.65	0.15	152	303	455
0.5	0.66	0.16	133	265	398
0.5	0.67	0.17	117	233	350
0.5	0.68	0.18	104	207	311
0.5	0.69	0.19	93	185	278
0.5	0.70	0.20	83	165	248

3.4. Statistical framework

The ACTION trial will use a superiority testing framework, with the active treatment arm (azithromycin) compared against placebo.

The analysis will account for interim analyses using a group-sequential testing framework using the Lan-DeMets spending function approach with an O'Brien and Fleming boundary [2].

3.5. Statistical interim analyses and stopping guidance

Due to uncertainty in the possible effects of treatment and the urgency of identifying effective therapies for COVID-19, sequential randomized controlled trials with interim monitoring are expected to provide robust results faster than standard fixed parallel trials without interim monitoring [3]. In the absence of a significant difference between randomized arms, the study team believes that there is value in continuing the randomized comparison of Azithromycin versus placebo to full enrollment (i.e., 455 participants) in order to obtain as much precision as possible and to provide maximal information to inform the field. We therefore have not recommended interim analyses for futility.

Interim analyses for efficacy

We have assumed 1 interim analysis at 50% of participants with 14-day endpoints measured using a Lan-DeMets spending approach with an O'brien-Fleming boundary. The actual alpha spending will be updated according to the actual interim analysis schedule, should trial investigators change the schedule in coordination with the trial's DSMC.

A summary of the interim analysis and nominal *P*-values is below, created by the R package gsDesign. We have also included a summary of standardized treatment effects along the efficacy bound, For additional details see this SAP's companion computational notebook: `ACTION-sample-size-power-symptom-free.Rmd / .html`.

```
## One-sided group sequential design with
## 80 % power and 2.5 % Type I Error.
##
##   Analysis  N   Z   Nominal p  Spend
##          1 228 2.96    0.0015 0.0015
##          2 456 1.97    0.0245 0.0235
##      Total                    0.0250
##
## ++ alpha spending:
##  Lan-DeMets O'Brien-Fleming approximation spending function with none = 1.
##
## Boundary crossing probabilities and expected sample size
## assume any cross stops the trial
##
## Upper boundary (power or Type I Error)
##           Analysis
##    Theta      1      2 Total  E{N}
##   0.0000 0.0015 0.0235 0.025 454.8
##   0.1316 0.1641 0.6359 0.800 417.9
```

We note the pre-specification of a flexible alpha spending function permits interim analysis at the discretion of the DSMC. We will update the efficacy bound according to the trial's actual interim analysis schedule.

The DSMC will review interim analyses and will make a recommendation to stop the trial for efficacy after each review. Early stopping for any reason will mandate final analysis at alpha of the full final prespecified value based on all available data (incorporating any alpha spending, as needed). In considering possible modifications to the study or termination of the study, the DSMC may consider other outcome measures than the composite hospitalization/death outcome measure. For example, the DSMC might make recommendations based on a high level of evidence for a difference between randomized arms in the proportion dying. A "high level of evidence" might be based on application of the O'Brien and Fleming stopping guideline to the death outcome.

3.6. Timing of the final analysis

The final analysis will take place when 14-day outcomes have been measured on all participants. If the trial is stopped for efficacy or for any other reason before the full sample, the final analysis will include all outcomes among patients who have been enrolled at the time the trial is stopped.

3.7. Timing of outcome assessments

Primary outcomes will be measured 14-days after enrollment. Additional visits will take place at 3, 7, and 21 days after enrollment. Protocol section 3.2 includes details about the trial profile.

4. Statistical Principles

4.1. Confidence intervals and *P*-values

The trial will report 95% confidence intervals and permutation *P*-values estimated from the primary analysis. The level of statistical significance will depend on the alpha spending function that accounts for the group sequential design (see section 3.5 for details).

4.2. Protocol deviations

The analysis population will be intention to treat, and will include all participants enrolled and randomized. Adherence to protocol will be determined during the 14-day assessment based on patient report of whether they took the medication delivered by the trial. At that time, the trial will assess whether patients have taken other medication during the study period, including azithromycin or other macrolides not provided by the trial. Patients who take azithromycin or other macrolide not provided by the study will be considered protocol deviations but will be included according to randomized group in the intention-to-treat analysis.

5. Trial Population

5.1. Screening data

We will report the number of patients screened and characteristics to the extent they are available to assess representativeness of the enrolled study population.

5.2. Eligibility

See Section 4 of the trial protocol for eligibility criteria.

5.3. Recruitment

See section 5.1 of the study protocol. We will report the number of participants screened, enrolled, randomized, and measured at the primary endpoint, along with reasons for exclusion at each step following CONSORT guidelines.

5.4. Withdrawal/follow-up

We will report the proportion of patients in both study arms who withdraw from the trial, along with reasons. We will report this information in the trial's CONSORT flow chart. The analysis will be intention-to-treat: patients will be analyzed according to the arm they are randomized.

5.5. Baseline patient characteristics

We will summarize patient characteristics by arm, including: age, sex, location (to level of US state), and clinical symptoms at the time of enrollment.

6. Analysis

6.1. Outcome definitions

Primary Outcome:
The primary outcome will be free of symptoms at 14 days after the electronic enrollment assessment. Symptoms include: cough, fever, myalgia, anosmia, shortness of breath (and related abilities such as ability to walk across a room or up a flight of stairs), fatigue, conjunctivitis, and orthostatic symptoms.

Secondary Outcomes:

- **Hospitalization and/or death by 14 days** after the electronic enrollment assessment.
- **Hospitalization/death by 21 days:** we will extend follow-up beyond the primary endpoint to capture any hospitalizations or death that progress more slowly and are not detected by 14 days.
- **Adverse events:** we will conduct an adverse event survey at Day 3 after treatment, including gastrointestinal side effects (nausea, vomiting, diarrhea, abdominal pain) and rash. We note that some COVID-19 patients report gastrointestinal symptoms, and this survey will provide data on whether azithromycin causes additional gastrointestinal effects beyond those symptoms.
- **Prevalence of positive swabs**: we will compare the prevalence of SARS-CoV-2 positive swabs at Day 3 after treatment in self-collected nasal, saliva, and rectal swabs in azithromycin compared to placebo-treated participants.
- **Viral load:** we will assess viral load by RT-PCR in self-collected nasal, saliva, and rectal swabs at Day 3 after treatment.
- **Mortality:** we will collect emergency contact/next of kin information during the baseline questionnaire. We will follow-up with the emergency contact if participants are lost to follow-up at the Day 14 and Day 21 questionnaires to assess mortality and hospitalization outcomes.
- **Genetic macrolide resistance determinants:** We will evaluate the prevalence of genetic macrolide resistance determinants ermB, mefA/E, and mphA by targeted PCR in rectal samples collected at Day 3 after treatment.
- **COVID-19 symptoms at 21 days:** we will ask patients about COVID-19 symptomology during each online questionnaire, including cough, fever, myalgia, anosmia, shortness of breath (and related abilities such as ability to walk across a room or up a flight of stairs), fatigue, conjunctivitis, and orthostatic symptoms.
- **Number of emergency room visits:** during each online questionnaire we will survey patients about any emergency room visits (with stays <24 hours) that occurred since their last survey.
- **Number of household members with COVID-19 (confirmed or symptomatic):** during each online questionnaire we will ask participants how many of their household members have symptoms of COVID-19 or confirmed disease.

6.2. Analysis methods

Primary Outcome:
The primary analysis will estimate the risk difference (RD) comparing the proportion of patients who are symptom-free at 14 days in the azithromycin versus placebo arms. We will compute *P*-

values for differences between arms using a permutation test with the prevalence difference between arms as the test statistic and 10,000 iterations. We will additionally estimate and report the prevalence ratio.

Determination of statistical significance for the primary and secondary analyses will follow the final alpha-spending function determined by the interim analyses (see section 3.5).

Secondary Outcomes:

- **Hospitalization/death by 21 days:** We will use a Kaplan-Meier approach to account for losses to follow-up. We will use Greenwood's formula to estimate the standard error of the log proportion in each arm will be used to determine the variance of the risk difference at 21 days.
- **Mortality by 21 days:** we will collect emergency contact/next of kin information during the baseline questionnaire. We will follow-up with the emergency contact if participants are lost to follow-up at the Day 14 and Day 21 questionnaires to assess mortality and hospitalization outcomes.
- **COVID-19 symptoms at 21 days:** we will ask patients about COVID-19 symptomology during each online questionnaire, including cough, fever, myalgia, anosmia, shortness of breath (and related abilities such as ability to walk across a room or up a flight of stairs), fatigue, conjunctivitis, and orthostatic symptoms.
- **COVID-19 (confirmed or symptomatic) among household members of enrolled patients:** during each online questionnaire we will ask participants how many of their household members have symptoms of COVID-19 or confirmed disease. This analysis will be conducted on an expanded dataset that includes an observation for each household member. To account for outcome correlation among individuals within households, we will estimate 95% confidence intervals for the risk difference and risk ratio with a clustered, non-parametric bootstrap that resamples households with replacement (10,000 iterations).

Outcomes with a prevalent measure at Day 3:

We will estimate the difference in prevalence between arms and the 95% confidence interval for the difference. We will compute *P*-values for differences between arms using a permutation test with the prevalence difference between arms as the test statistic and 10,000 iterations.

- **Prevalence of positive swabs**: we will compare the prevalence of SARS-CoV-2 positive swabs at Day 3 after treatment in self-collected nasal, saliva, and rectal swabs in azithromycin compared to placebo-treated participants.
- **Genetic macrolide resistance determinants:** We will evaluate the prevalence of genetic macrolide resistance determinants ermB, mefA/E, and mphA by targeted PCR in rectal samples collected at Day 3 after treatment.
- **Adverse events:** we will conduct an adverse event survey at Day 3 after treatment, including gastrointestinal side effects (nausea, vomiting, diarrhea, abdominal pain) and rash. We note that some COVID-19 patients report gastrointestinal symptoms, and this survey will provide data on whether azithromycin causes additional gastrointestinal effects beyond those symptoms. We also note that adverse events will be cumulative through Day 3 and measure cumulative incidence (risk) not prevalence.

Outcomes with a continuous measure at Day 3:

We will compare log transformed values of relative read and number of reads using a t-test. We will compute *P*-values for differences between arms using a permutation test with the difference between arms as the test statistic and 10,000 iterations.

- **Viral load:** we will assess viral load by RT-PCR in self-collected nasal, saliva, and rectal swabs at Day 3 after treatment.

Count outcomes:

- **Number of emergency room visits:** during each online questionnaire we will survey patients about any emergency room visits (with stays <24 hours) that occurred since their last survey. We will model the total number of emergency room visits for each participant through 14 days as a count outcome using a negative binomial regression model with a single indicator for treatment arm and an offset for the number of days of follow-up completed by each participant. The exponentiated coefficient on the indicator will estimate the incidence rate ratio. We will compute a *P*-value for difference between arms using a permutation test with the coefficient from the model as the test statistic and 10,000 iterations.

6.3. Missing data

If outcome data are missing for >15% of participants we will report results from sensitivity analyses for missing data [4,5]. We selected 15% missing because trials with >20% missing values are thought to be a concern for bias [6]. The primary analysis will account for censoring but will assume outcomes are missing completely at random. We will relax this assumption slightly, assuming that outcomes are missing at random (MAR) with an inverse-probability weighted (IPW) estimator using cumulative incidence through 14 days as the binary endpoint. We will model the probability of censoring as a function of baseline patient characteristics with a logistic regression model, and will use the inverse of the predicted probabilities from the model to re-weight the analysis population to reflect the full study population at the time of enrollment, using a doubly-robust targeted maximum likelihood estimator that also includes an outcome model using the same covariates [7]. Baseline characteristics used in the censoring model will include age, sex, race/ethnicity, zip code, comorbidities, concomitant medications, symptoms at enrollment, insurance status, alcohol use, tobacco/e-cigarette/marijuana use, and household composition. We will additionally consider a range of scenarios assuming the outcomes are missing not at random (MNAR), *i.e.* systematically, using a pattern mixture model approach described by Little et al. [4], whereby we model cumulative incidence through 14 days with a linear probability model as a function of covariates listed for the IPW estimator; missing outcomes will then be imputed using the model fit to predict , adding a shift parameter Δ that we vary across a range of values.

6.4. Additional analyses

Pre-specified Subgroup Analyses

In secondary analyses, we will estimate differences separately by the following subgroups:

- Age >60 versus ≤ 60
- "High risk" versus low risk patients, with "high risk" defined as Age >60 and reported hypertension, cardiovascular disease, diabetes, or obstructive or restrictive lung disease at enrollment.
- Presence of COVID-19 symptoms at enrollment versus asymptomatic at enrollment.

6.5. Harms

Section 10 of the trial protocol includes details of adverse event monitoring and reporting. All serious adverse events including deaths will be recorded for the period between the date the first dose of azithromycin or placebo and the date of the primary endpoint, 14 days later, and through the final follow-up at 21 days.

6.6. Statistical software

Analyses will be conducted using R version 4.0 or later. We will strive to make all replication files publicly available along with the trial results through the Proctor Foundation's GitHub page and other public repositories such as the Open Science Framework and Zenodo.

7. References

1. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318: 2337–2343. doi:10.1001/jama.2017.18556

2. DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. Stat Med. 1994;13: 1341–1352; discussion 1353-1356. doi:10.1002/sim.4780131308

3. Cooper BS, Boni MF, Pan-ngum W, Day NPJ, Horby PW, Olliaro P, et al. Evaluating clinical trial designs for investigational treatments of Ebola virus disease. PLoS Med. 2015;12: e1001815. doi:10.1371/journal.pmed.1001815

4. National Research Council (U.S.), National Research Council (U.S.), National Academies Press (U.S.), editors. The prevention and treatment of missing data in clinical trials. Washington, D.C: National Academies Press; 2010.

5. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367: 1355–1360. doi:10.1056/NEJMsr1203730

6. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet. 2002;359: 781–785. doi:10.1016/S0140-6736(02)07882-0

7. Gruber S, van der Laan M. tmle: An R Package for Targeted Maximum Likelihood Estimation. J Stat Softw. 2012;51: 1–35.

8. Revision history

Version	Date	Summary of Changes, Justification, and Timing vis-à-vis key trial events (enrollment completion, interim analyses, unmasking, etc)
1	2020-03-28	• First draft
2	2020-03-28	• Updated administrative information
3	2020-04-03	• Provided additional detail in the interim monitoring section.
4	2020-05-05	• Update to reflect discussions during the trial's first DSMC meeting on April 8, 2020, as well as input from the FDA through the IND approval process, including: • Add secondary outcomes beyond viral load at day 3, including analysis details for each • Add pre-specified subgroup analysis (previously none) • Remove the futility bound for interim analyses • Update the interim analysis alpha spending function to Lan-DeMets (harmonization with other COVID-19 therapeutic trials) • Update the primary analysis to be based on a Kaplan-Meier estimator rather than cumulative proportions to explicitly account for censoring. • Add language regarding a possible expansion of sample size to target a smaller detectable effect (RR = 0.8 rather than RR = 0.6) if large-scale enrollment is possible. • Made risk difference the primary comparison with risk ratio (RR) an additional contrast we plan to report. • Include details of masked sample-size re-assessment committee, and specific timing for masked sample size re-assessment. • Add details of possible sensitivity analyses for missing outcomes under MAR and MNAR
5	2020-05-29	• Revised the interim sample size evaluation process and group who will be involved in the decision making process. (version 5 included in the Appendix)
6	2020-12-15	• Changed the trial's primary outcome from hospitalization/death by 14-d to symptom free at 14-d • Updated the sample size calculation for the new outcome, revising the trial's planned enrollment to N=455. • Updated the interim analysis plan to propose a single interim efficacy analysis when 50% of participants have reached the primary endpoint. • Simplified the missing data section with demotion of hospitalization/death to a secondary outcome, since the interim analyses will no longer include that outcome.

9. Appendix

Azithromycin for COVID-19 Treatment In Outpatients Nationwide (ACTION)

Statistical Analysis Plan

1. Administrative Information

Trial registration:	ClinicalTrials.gov NCT04332107
Funder:	Bill & Melinda Gates Foundation (grant # INV-017026)
SAP Version:	Version 5, Updated 2020-05-29 A revision history for this document is included at the end.
Protocol Version:	Refers to Protocol Version 1.17
Contributors:	Statisticians responsible: Benjamin F. Arnold [1] , Travis C. Porco [1] Principal Investigators: Catherine E. Oldenburg [1], Thuy Doan[1]

[1] Francis I. Proctor Foundation for Research in Ophthalmology, University of California, San Francisco, USA

SAP Table of Contents

The content of this Statistical Analysis Plan meets the requirements stated by the US Food and Drug Administration and conforms to the American Statistical Association's Ethical Guidelines. This SAP was organized following guidelines proposed in:

Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA*. 2017;318: 2337–2343. PMID: 29260229 [1]

The companion computational notebook with underlying sample size and group sequential design calculations presented herein is entitled: `ACTION-sample-size-power.Rmd / .html`; it is saved in the same directory as this document.

2. Introduction

2.1. Background and rationale

Identification of a safe, effective treatment for individuals with mild or moderate COVID-19 that prevents disease progression and reduces hospitalization would reduce the burden on the health system. High dose hydroxychloroquine is being evaluated for SARS-CoV-2 prevention and COVID-19 disease treatment, but has a high risk of a number of potentially severe adverse events. Recent evidence has indicated that the broad-spectrum macrolide azithromycin may have some activity against coronaviruses. A large community randomized trial in Niger demonstrated reduced viral load among children with commensal coronaviruses (prior to the emergence of SARS-CoV-2) in communities receiving biannual mass azithromycin distribution compared to placebo (submitted). In SARS-CoV-2 patients in France, the addition of azithromycin to a hydroxychloroquine regimen appeared to decrease SARS-CoV-2 positivity by PCR compared to hydroxychloroquine-only and control patients. Azithromycin is generally well-tolerated and may be an attractive option for treating patients with mild disease.

2.2. Objectives

1: To determine if a single oral dose of azithromycin is effective for reducing progression to hospitalization in patients with a positive SARS-CoV-2 PCR test with mild or moderate disease.

3. Study Methods

3.1. Trial design

The trial will be a scalable, telemedicine-based parallel group simple trial to evaluate the efficacy of azithromycin for the prevention of COVID-19 disease progression and hospitalization.

3.2. Randomization

Participants will be randomized 2:1 to either receive a dose of oral azithromycin or oral placebo. Patients will be twice as likely to receive azithromycin than placebo. The randomization will not be stratified or blocked in any way. We found that even with permuted block sizes and ≥6 labels (4 azithromycin, 2 placebo), a 2:1 allocation could lead to unmasking through longer runs of active treatment labels.

3.3. Sample size

The trial's initial target enrollment will be on the order of 2,300 participants, which will be sufficient to detect a reduction of proportion of patients who are hospitalized from 10% to 6% (4 percentage point reduction) allowing for 7 interim analyses (details below). The trial's final enrollment could be as high as 10,000 participants, depending on how the trial evolves.

Due to unknown true incidence, probability of hospitalization, and enrollment potential, we propose an adaptive sample size that is re-evaluated at pre-specified junctures by a masked group of investigators that will include a masked statistician (TC Porco), only provided incidence and enrollment data (masked to study arm) for sample size re-estimation, along with study PIs (Drs. Oldenburg, Doan) and Dr. Lietman. The sample size evaluation will be based on primary outcome rates estimated across all three arms (pooled), and will be completed after 30% (n=690) and 50% (n=1,150) of participants reach the primary endpoint. The masked investigator

group will consider recommending change to sample size based not only on statistical considerations but also on logistical considerations such as the pace of enrollment and the course of the epidemic. The masked investigators will make a final recommendation to the trial's Data and Safety Monitoring Committee and institutional review board. Masked interim sample size re-estimation is not thought to strongly influence Type I error [2].

With 2,268 participants, the trial will have 80% power to detect a relative reduction of 40% in hospitalization or death under the following assumptions:

- 2:1 allocation ratio of azithromycin : placebo
- 10% hospitalization under placebo over 14-day period (binomial outcome, 14-d cumulative incidence)
- 10% of patients seek azithromycin therapy outside of the trial.
    - This non-compliance will attenuate the ITT estimates slightly by making the outcome rate more similar between the two groups. For example, if we assume 10% hospitalization in the control arm and that treatment reduces the percentage to 6% (a risk difference of 4%), then this non-compliance means that the actual outcome prevalence in the placebo arm will be (0.9)(10%)+(0.10)(6%)=9.6%
- 15% loss to follow-up
- 2-sided alpha of 0.05, power = 0.8.
- Seven interim analyses at 30%, 40%, …, 90% of outcomes collected with stopping guideline determined using the Lan-DeMets spending function approach with an O'Brien and Fleming boundary.

The table below shows sample sizes under five effect size scenarios, assuming azithromycin reduces hospitalization or death by between 1% (RR = 0.9) and 5% (RR = 0.5) compared to placebo.

Sample size required for a fixed design allowing for 15% loss to follow-up. Calculations assume 2-sided alpha of 0.05 and 80% power.

p placebo	p azithro	RD (ITT)*	RR (ITT)*	N placebo	N azithro	N Total
0.099	0.09	-0.009	0.909	14,221	28,441	42,662
0.098	0.08	-0.018	0.816	3,346	6,691	10,037
0.097	0.07	-0.027	0.722	1,392	2,784	4,176
0.096	0.06	-0.036	0.625	729	1,457	2,186
0.095	0.05	-0.045	0.526	431	861	1,292

* Risk difference (RD) and risk ratio (RR) under an ITT analysis assuming that 10% of all participants seek azithromycin from sources outside of the trial. This noncompliance has been included in the assumptions of the probability of hospitalization under placebo, from a base of 0.1.

Note, these estimates do not account for interim analyses in a group-sequential analysis. Sequential randomized controlled trials have been demonstrated to result in overall better patient outcomes during fast-moving epidemics because they reach answers for efficacy more quickly than a trial without interim analyses [3].

Using the scenario of a 4% reduction in hospitalization due to placebo, we estimated the maximum sample size under a group-sequential design. Under the assumptions in Section 3.5

(below), the maximum sample size would be **2,268**, a 3.5% increase over a design with no interim monitoring due to alpha spending at 7 interim analyses.

We note that ACTION is designed as a scalable trial. If pre-specified, interim sample size calculations or external factors suggest that increasing the sample size would be desirable, then the masked investigator group that evaluates interim sample size assessments at 30% and 50% enrollment will make a recommendation to the trial's Data and Safety Monitoring Committee, which together will decide if such expansion is desirable. An expansion to approximately 10,000 participants would allow for the detection of smaller but still clinically relevant effects (RR=0.8) given the severity of hospitalization or death.

3.4. Statistical framework

The ACTION trial will use a superiority testing framework, with the active treatment arm (azithromycin) compared against placebo.

The analysis will account for interim analyses using a group-sequential testing framework using the Lan-DeMets spending function approach with an O'Brien and Fleming boundary [4].

3.5. Statistical interim analyses and stopping guidance

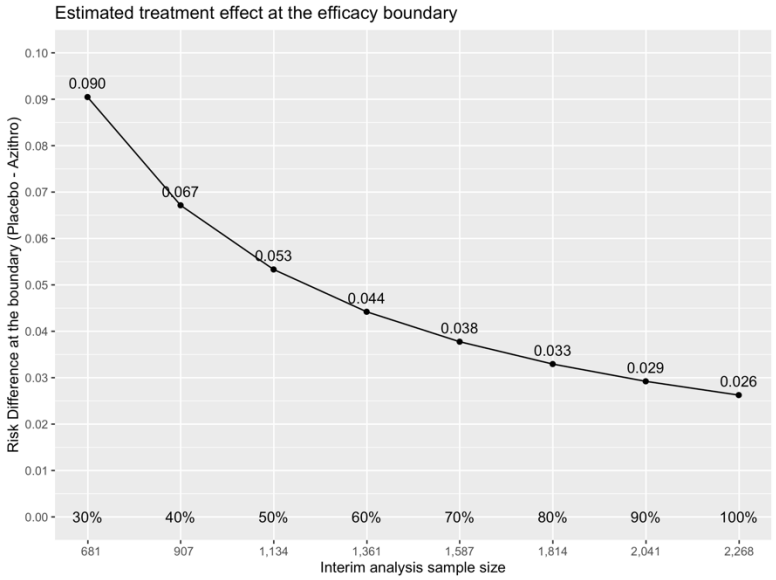
Due to uncertainty in the possible effects of treatment and the urgency of identifying effective therapies for COVID-19, sequential randomized controlled trials with interim monitoring are expected to provide robust results faster than standard fixed parallel trials without interim monitoring [3]. In the absence of a significant difference between randomized arms, the study team believes that there is value in continuing the randomized comparison of Azithromycin versus placebo to full enrollment (i.e., 2300 participants) in order to obtain as much precision as possible and to provide maximal information to inform the field. We therefore have not recommended interim analyses for futility.

Interim analyses for efficacy
The efficacy bound was derived using a Lan-DeMets spending approach with an O'brien-Fleming boundary. At this time, we have assumed 7 interim analyses at 30%, 40%, …, 90% of participants with 14-day endpoints measured. This is an approximation as the trial is currently considering weekly interim analyses. The actual alpha spending will be updated according to the actual interim analysis schedule.

A summary of the interim analyses and nominal *P*-values is below, created by the R package gsDesign. We have also included a summary of standardized treatment effects along the efficacy bound, For additional details see this SAP's companion computational notebook: `ACTION-sample-size-power.Rmd / .html`.

```
## One-sided group sequential design with
## 80 % power and 2.5 % Type I Error.
##
##   Analysis    N    Z   Nominal p  Spend
##          1  681 3.93    0.00004 0.00004
##          2  907 3.37    0.0004  0.0004
##          3 1134 2.99    0.0014  0.0011
##          4 1361 2.71    0.0033  0.0023
##          5 1587 2.50    0.0061  0.0036
##          6 1814 2.34    0.0098  0.0048
##          7 2041 2.20    0.0140  0.0059
##          8 2268 2.08    0.0187  0.0069
##      Total                       0.0250
##
## ++ alpha spending:
##  Lan-DeMets O'brien-Fleming approximation spending function.
##
## Boundary crossing probabilities and expected sample size
## assume any cross stops the trial
##
## Upper boundary (power or Type I Error)
##           Analysis
##    Theta      1      2      3      4      5      6      7      8 Total   E{N}
##   0.0000 0.000 0.0004 0.0011 0.0023 0.0036 0.0048 0.0059 0.0069 0.025 2257.1
##   0.0599 0.009 0.0509 0.1099 0.1471 0.1524 0.1359 0.1104 0.0844 0.800 1735.2
```

The figure below shows the difference in the proportion of patients hospitalized (placebo minus azithromycin) at each of the interim analyses along the efficacy bound based on the alpha spending function above. It shows, for example, that with 681 patients enrolled at the first interim analysis, hospitalization would need to be 9 percentage points higher among patients in the placebo arm compared with the azithromycin arm to stop the trial for efficacy.



We note the pre-specification of a flexible alpha spending function permits interim analysis at the discretion of the DSMC. We will update the efficacy bound according to the trial's actual interim analysis schedule.

The DSMC will review interim analyses and will make a recommendation to stop the trial for efficacy after each review. Early stopping for any reason will mandate final analysis at alpha of the full final prespecified value based on all available data (incorporating any alpha spending, as

needed). In considering possible modifications to the study or termination of the study, the DSMC may consider other outcome measures than the composite hospitalization/death outcome measure. For example, the DSMC might make recommendations based on a high level of evidence for a difference between randomized arms in the proportion dying. A "high level of evidence" might be based on application of the O'Brien and Fleming stopping guideline to the death outcome.

3.6. Timing of the final analysis

The final analysis will take place when 14-day outcomes have been measured on all participants. If the trial is stopped for efficacy or futility before the full sample, the final analysis will include all outcomes among patients who have been enrolled at the time the trial is stopped.

3.7. Timing of outcome assessments

Primary outcomes will be measured 14-days after enrollment. Additional visits will take place at 3, 7, and 21 days after enrollment. Protocol section 3.2 includes details about the trial profile.

4. Statistical Principles

4.1. Confidence intervals and *P*-values

The trial will report 95% confidence intervals and permutation *P*-values estimated from the primary analysis. The level of statistical significance will depend on the alpha spending function that accounts for the group sequential design (see section 3.5 for details).

4.2. Protocol deviations

The analysis population will be intention to treat, and will include all participants enrolled and randomized. Adherence to protocol will be determined during the 14-day assessment based on patient report of whether they took the medication delivered by the trial. At that time, the trial will assess whether patients have taken other medication during the study period, including azithromycin or other macrolides not provided by the trial. Patients who take azithromycin or other macrolide not provided by the study will be considered protocol deviations but will be included according to randomized group in the intention-to-treat analysis.

5. Trial Population

5.1. Screening data

We will report the number of patients screened and characteristics to the extent they are available to assess representativeness of the enrolled study population.

5.2. Eligibility

See Section 4 of the trial protocol for eligibility criteria.

5.3. Recruitment

See section 5.1 of the study protocol. We will report the number of participants screened, enrolled, randomized, and measured at the primary endpoint, along with reasons for exclusion at each step following CONSORT guidelines.

5.4. Withdrawal/follow-up

We will report the proportion of patients in both study arms who withdraw from the trial, along with reasons. We will report this information in the trial's CONSORT flow chart. The analysis will be intention-to-treat: patients will be analyzed according to the arm they are randomized.

5.5. Baseline patient characteristics

We will summarize patient characteristics by arm, including: age, sex, location (to level of US state), and clinical symptoms at the time of enrollment.

6. Analysis

6.1. Outcome definitions

Primary Outcome:
The primary outcome will be hospitalization and/or death at 14 days after the electronic enrollment assessment.

Secondary Outcomes:

- **Hospitalization/death by 21 days:** we will extend follow-up beyond the primary endpoint to capture any hospitalizations or death that progress more slowly and are not detected by 14 days.
- **Adverse events:** we will conduct an adverse event survey at Day 3 after treatment, including gastrointestinal side effects (nausea, vomiting, diarrhea, abdominal pain) and rash. We note that some COVID-19 patients report gastrointestinal symptoms, and this survey will provide data on whether azithromycin causes additional gastrointestinal effects beyond those symptoms.
- **Prevalence of positive swabs**: we will compare the prevalence of SARS-CoV-2 positive swabs at Day 3 after treatment in self-collected nasal, saliva, and rectal swabs in azithromycin compared to placebo-treated participants.
- **Viral load:** we will assess viral load by RT-PCR in self-collected nasal, saliva, and rectal swabs at Day 3 after treatment.
- **Mortality:** we will collect emergency contact/next of kin information during the baseline questionnaire. We will follow-up with the emergency contact if participants are lost to follow-up at the Day 14 and Day 21 questionnaires to assess mortality and hospitalization outcomes.

- **Genetic macrolide resistance determinants:** We will evaluate the prevalence of genetic macrolide resistance determinants ermB, mefA/E, and mphA by targeted PCR in rectal samples collected at Day 3 after treatment.
- **COVID-19 symptoms:** we will ask patients about COVID-19 symptomology during each online questionnaire, including cough, fever, myalgia, anosmia, shortness of breath (and related abilities such as ability to walk across a room or up a flight of stairs), fatigue, conjunctivitis, and orthostatic symptoms.
- **Number of emergency room visits:** during each online questionnaire we will survey patients about any emergency room visits (with stays <24 hours) that occurred since their last survey.
- **Number of household members with COVID-19 (confirmed or symptomatic):** during each online questionnaire we will ask participants how many of their household members have symptoms of COVID-19 or confirmed disease.

6.2. Analysis methods

Primary Outcome:
The primary analysis will estimate the risk difference (RD) comparing cumulative incidence of hospitalization or death by 14 days in the azithromycin versus placebo arms. Kaplan-Meier approach to account for losses to follow-up. We will use Greenwood's formula to estimate the standard error of the log proportion in each arm will be used to determine the variance of the risk difference, and will report 95% confidence intervals adjusted for interim analyses. We will compute *P*-values for differences between arms using a permutation test with the risk difference between arms as the test statistic and 10,000 iterations. We will additionally estimate and report the risk ratio.

Determination of statistical significance for the primary and secondary analyses will follow the final alpha-spending function determined by the group sequential design (see section 3.5).

Secondary Outcomes:

Outcomes with analyses that follow the same approach as the primary outcome:

Analyses for the following secondary endpoints will follow the same analysis methods as the for the primary outcome, a Kaplan-Meier analysis to estimate cumulative incidence, and *P*-values for differences between arms using a permutation test with the risk difference between arms as the test statistic and 10,000 iterations.

- **Hospitalization/death by 21 days:** we will extend follow-up beyond the primary endpoint to capture any hospitalizations or death that progress more slowly and are not detected by 14 days.
- **Mortality by 14 and 21 days:** we will collect emergency contact/next of kin information during the baseline questionnaire. We will follow-up with the emergency contact if participants are lost to follow-up at the Day 14 and Day 21 questionnaires to assess mortality and hospitalization outcomes.
- **COVID-19 symptoms:** we will ask patients about COVID-19 symptomology during each online questionnaire, including cough, fever, myalgia, anosmia, shortness of breath (and

related abilities such as ability to walk across a room or up a flight of stairs), fatigue, conjunctivitis, and orthostatic symptoms.
- **COVID-19 (confirmed or symptomatic) among household members of enrolled patients:** during each online questionnaire we will ask participants how many of their household members have symptoms of COVID-19 or confirmed disease. This analysis will be conducted on an expanded dataset that includes an observation for each household member. To account for outcome correlation among individuals within households, we will estimate 95% confidence intervals for the risk difference and risk ratio with a clustered, non-parametric bootstrap that resamples households with replacement (10,000 iterations).

Outcomes with a prevalent measure at Day 3:

We will estimate the difference in prevalence between arms and the 95% confidence interval for the difference. We will compute *P*-values for differences between arms using a permutation test with the prevalence difference between arms as the test statistic and 10,000 iterations.

- **Prevalence of positive swabs**: we will compare the prevalence of SARS-CoV-2 positive swabs at Day 3 after treatment in self-collected nasal, saliva, and rectal swabs in azithromycin compared to placebo-treated participants.
- **Genetic macrolide resistance determinants:** We will evaluate the prevalence of genetic macrolide resistance determinants ermB, mefA/E, and mphA by targeted PCR in rectal samples collected at Day 3 after treatment.
- **Adverse events:** we will conduct an adverse event survey at Day 3 after treatment, including gastrointestinal side effects (nausea, vomiting, diarrhea, abdominal pain) and rash. We note that some COVID-19 patients report gastrointestinal symptoms, and this survey will provide data on whether azithromycin causes additional gastrointestinal effects beyond those symptoms. We also note that adverse events will be cumulative through Day 3 and measure cumulative incidence (risk) not prevalence.

Outcomes with a continuous measure at Day 3:

We will compare log transformed values of relative read and number of reads using a t-test. We will compute *P*-values for differences between arms using a permutation test with the difference between arms as the test statistic and 10,000 iterations.

- **Viral load:** we will assess viral load by RT-PCR in self-collected nasal, saliva, and rectal swabs at Day 3 after treatment.

Count outcomes:

- **Number of emergency room visits:** during each online questionnaire we will survey patients about any emergency room visits (with stays <24 hours) that occurred since their last survey. We will model the total number of emergency room visits for each participant through 14 days as a count outcome using a negative binomial regression model with a single indicator for treatment arm and an offset for the number of days of follow-up completed by each participant. The exponentiated coefficient on the indicator will estimate the incidence rate ratio. We will compute a *P*-value for difference between arms using a permutation test with the coefficient from the model as the test statistic and 10,000 iterations.

6.3. Missing data

If we find there is significant delay in ascertainment of the primary outcome during the trial, it could bias interim analyses. If, in the conduct of the trial, we find that outcome ascertainment is delayed for a significant number of participants, then we will consider adjusting our interim analysis Kaplan-Meier estimator using an inverse probability of censoring weighted approach, which weights the failure distribution by a Kaplan-Meier estimate of the censoring distribution [5].

If outcome data are missing for >15% of participants we will report results from sensitivity analyses for missing data [6,7]. We selected 15% missing because trials with >20% missing values are thought to be a concern for bias [8]. The primary analysis will account for censoring but will assume outcomes are missing completely at random. We will relax this assumption slightly, assuming that outcomes are missing at random (MAR) with an inverse-probability weighted (IPW) estimator using cumulative incidence through 14 days as the binary endpoint. We will model the probability of censoring as a function of baseline patient characteristics with a logistic regression model, and will use the inverse of the predicted probabilities from the model to re-weight the analysis population to reflect the full study population at the time of enrollment, using a doubly-robust targeted maximum likelihood estimator that also includes an outcome model using the same covariates [9]. Baseline characteristics used in the censoring model will include age, sex, race/ethnicity, zip code, comorbidities, concomitant medications, symptoms at enrollment, insurance status, alcohol use, tobacco/e-cigarette/marijuana use, and household composition. We will additionally consider a range of scenarios assuming the outcomes are missing not at random (MNAR), *i.e.* systematically, using a pattern mixture model approach described by Little et al. [6], whereby we model cumulative incidence through 14 days with a linear probability model as a function of covariates listed for the IPW estimator; missing outcomes will then be imputed using the model fit to predict , adding a shift parameter Δ that we vary across a range of values.

6.4. Additional analyses

Pre-specified Subgroup Analyses

In secondary analyses, we will estimate differences separately by the following subgroups:

- Age >60 versus ≤ 60
- “High risk” versus low risk patients, with “high risk” defined as Age >60 and reported hypertension, cardiovascular disease, diabetes, or obstructive or restrictive lung disease at enrollment.
- Presence of COVID-19 symptoms at enrollment versus asymptomatic at enrollment.

6.5. Harms

Section 10 of the trial protocol includes details of adverse event monitoring and reporting. All serious adverse events including deaths will be recorded for the period between the date the

first dose of azithromycin or placebo and the date of the primary endpoint, 14 days later, and through the final follow-up at 21 days.

6.6. Statistical software

Analyses will be conducted using R version 3.6 or later. We will strive to make all replication files publicly available along with the trial results through the Proctor Foundation's GitHub page and other public repositories such as the Open Science Framework and Zenodo.

7. References

1. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318: 2337–2343. doi:10.1001/jama.2017.18556

2. Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. Boca Raton: Chapman & Hall/CRC; 2000.

3. Cooper BS, Boni MF, Pan-ngum W, Day NPJ, Horby PW, Olliaro P, et al. Evaluating clinical trial designs for investigational treatments of Ebola virus disease. PLoS Med. 2015;12: e1001815. doi:10.1371/journal.pmed.1001815

4. DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. Stat Med. 1994;13: 1341–1352; discussion 1353-1356. doi:10.1002/sim.4780131308

5. Zhao H, Tsiatis A. A consistent estimator for the distribution of quality adjusted survival time. Biometrika. 1997;84: 339–348. doi:10.1093/biomet/84.2.339

6. National Research Council (U.S.), National Research Council (U.S.), National Academies Press (U.S.), editors. The prevention and treatment of missing data in clinical trials. Washington, D.C: National Academies Press; 2010.

7. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367: 1355–1360. doi:10.1056/NEJMsr1203730

8. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet. 2002;359: 781–785. doi:10.1016/S0140-6736(02)07882-0

9. Gruber S, van der Laan M. tmle: An R Package for Targeted Maximum Likelihood Estimation. J Stat Softw. 2012;51: 1–35.

8. Revision history

Version	Date	Summary of Changes, Justification, and Timing vis-à-vis key trial events (enrollment completion, interim analyses, unmasking, etc)
1	2020-03-28	• First draft
2	2020-03-28	• Updated administrative information
3	2020-04-03	• Provided additional detail in the interim monitoring section.
4	2020-05-05	• Update to reflect discussions during the trial's first DSMC meeting on April 8, 2020, as well as input from the FDA through the IND approval process, including: • Add secondary outcomes beyond viral load at day 3, including analysis details for each • Add pre-specified subgroup analysis (previously none) • Remove the futility bound for interim analyses • Update the interim analysis alpha spending function to Lan-DeMets (harmonization with other COVID-19 therapeutic trials) • Update the primary analysis to be based on a Kaplan-Meier estimator rather than cumulative proportions to explicitly account for censoring. • Add language regarding a possible expansion of sample size to target a smaller detectable effect (RR = 0.8 rather than RR = 0.6) if large-scale enrollment is possible. • Made risk difference the primary comparison with risk ratio (RR) an additional contrast we plan to report. • Include details of masked sample-size re-assessment committee, and specific timing for masked sample size re-assessment. • Add details of possible sensitivity analyses for missing outcomes under MAR and MNAR
5	2020-05-29	• Revised the interim sample size evaluation process and group who will be involved in the decision making process.