Statistical Analysis Plan

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

Table of Contents

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

Protocol

Repurposed Approved Therapies for Outpatient Treatment of Patients with Early-Onset COVID-19 and Mild Symptoms

Protocol Number:	TOGETHER Trial Master Protocol V4.0
(Version Date)	8-Feb-2022
Name of Test Drug:	Refer to the Master Protocol Appendix 1
Phase:	3
Methodology:	A placebo-controlled adaptive randomized platform trial
Sponsor:	McMaster University, Hamilton, Ontario
Principal Investigators:	Dr. Edward Mills Email: millsej@mcmaster.ca Dr. Gilmar Reis Email: administrador@cardresearch.org
Document Date:	8-Feb-2022
Document Version:	TOGETHER MP SAP Version 3.0

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

SIGNATURE PAGE

Protocol Title:	Master Clinical Trial Protocol: Repurposed Approved Therapies for Outpatient Treatment of Patients with Early-Onset COVID-19 and Mild Symptoms
Principal Investigators:	Edward Mills Email: millsej@mcmaster.ca Gilmar Reis, CardResearch Email: administrador@cardresearch.org
Protocol Number:	Version 4.0
Document Date/Version:	8-Feb-2022
Steve Kanters, PhD Scientific Director, RainCity Analytics	Signature: DocuSigned by: Steve Kanters DBCFA4A0F5C74CC... Date: 2/14/2022
Eve Limbrick-Oldfield, PhD Associate Director, RainCity Analytics	Signature: DocuSigned by: Eve Limbrick-Oldfield E5EDF9C8518F47F... Date: 2/14/2022

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

Sponsor Approval

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidance and guidelines.

I have discussed any questions I have regarding the contents of this document with the biostatistical author.

I also understand that any subsequent changes to the planned statistical analyses, as described herein, may have a regulatory impact and/or result in timeline adjustments. All changes to the planned analyses will be described in the clinical study report (CSR).

Principal Investigator Signatory:
Edward Mills

Email: millsej@mcmaster.ca

Signature: DocuSigned by: Ed Mills 091855DB328C4C4...

Date: 2/14/2022

Principal Investigator Signatory:
Gilmar Reis

Email: administrador@cardresearch.org

Signature: DocuSigned by: Gilmar Reis 6165DB6742E14C5...

Date: 2/14/2022

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

ABBREVIATIONS

Abbreviation	Definition
ATC	Anatomic Therapeutic Class
BUGS	Bayesian inference using Gibbs Sampling
CER	Control Event Rate
COVID-19	Coronavirus disease 2019
CRF	Case report form
CSR	Clinical Study Report
ICH	International Council Harmonisation
IP	Investigational Product
ITT	Intention-To-Treat
JAGS	Just Another Gibbs Sampler
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
PP	Per-Protocol
PROMIS	Patient-Reported Outcomes Management Information System
RRR	Relative Risk Reduction
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SAP	Statistical Analysis Plan
SOC	System/Organ/Class
WHO	World Health Organization

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

1. INTRODUCTION AND OBJECTIVES OF ANALYSIS

1.1. Introduction

The TOGETHER master protocol represents an adaptive platform trial for treating persons with early severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection at high risk of disease progression who do not require hospital admission. The TOGETHER trial will start with a placebo as a control in the clinical evaluation of repurposed investigational products (IPs) to treat SARS-CoV-2. Other affordable candidate drug regimens that can be repurposed for coronavirus disease 2019 (COVID-19) may be considered and incorporated into this trial as an additional arm (s). If an intervention is shown to be effective, this design will allow the replacement of the placebo group with the effective intervention as the comparator. To ensure timely results the master protocol allows for data from trials that follow a similar protocol as TOGETHER to be included in the analyses.

1.2. Objectives of Statistical Analysis

This statistical analysis plan (SAP) is designed to outline the methods to be used in the analysis of study data to answer the study objective(s). We provided the populations for analysis, data handling rules, statistical methods, and data presentation formats. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of the clinical study report (CSR) for this trial.

Sponsor:	McMaster University		
Protocol:	TOGETHER MP		
Document Version No.:	3.0	**Document Date:**	8-Feb-2022

2. STUDY DESIGN

2.1. Synopsis of Study Design

TOGETHER trial is an international multicenter adaptive randomized platform trial for the early treatment of SARS-CoV-2 infection in high-risk adults not requiring hospital admission. The TOGETHER Trial will start with a placebo as a control in the clinical evaluation of three IPs. This trial is designed as a platform trial design that can add new arms onto the trial under standardized eligibility criteria, outcomes, and measurements, as the other experimental interventions. Other affordable candidate drug regimens that can be repurposed for COVID-19 may also be considered for this trial. The decision to add new therapeutic strategies will be based on external findings with local stakeholders' consultations.

2.2. Randomization Methodology

Please refer to the master protocol for the randomization methodology.

2.3. Stopping Rules and Unblinding

Stopping rules are outlined in section 4.3. In the case of potential recruitment challenges, it is important to reach statistical conclusions about the experimental treatments as fast as possible; therefore, continual Bayesian learning methods have been prepared and simulated to prepare for potential protocol changes. These simulation results are described in section 4.3.

2.4. Study Procedures

Please refer to the study protocol for the study procedures.

2.5 Efficacy and Safety Variables

2.5.1 Efficacy Variables

Primary endpoint:

- Emergency room visit due to the clinical worsening of COVID-19 (defined as participants remaining under observation for > 6 hours) within 28 days of randomization OR
- Hospitalization due to the progression of COVID-19 (defined as worsening of viral pneumonia) and/or complications within 28 days of randomization

Secondary endpoints:

- All cause mortality, including COVID-19 related mortality
- Reduction of viral load after randomization (D3 and D7)
- Number of days with respiratory symptoms
- Time to hospitalization/urgent care due to the progression of COVID-19
- Time to hospitalization for any cause
- Hospitalization for any cause
- Time to death
- Time to clinical improvement
- Symptoms as assessed by the WHO Clinical Worsening Scale
- Symptoms as assessed by the Wisconsin Upper Respiratory Symptom Survey (WURSS-21)
- Oxygen saturation ≤ 93% whilst under medical supervision
- Health-related quality of life as assessed by PROMIS global health scale (“Global-10”) scores (day 14 and day 28)
- Cognitive ability at day 28 (Telephone Interview for Cognitive Status)

- Adverse events
- Adverse reactions to the study medications
- Adherence with the IPs

Exploratory endpoints:

- Number of days spent in an intensive care unit
- Number of days on invasive mechanical ventilation
- Number of days of hospitalization
- Number of days of hospitalization in the ward
- Number of days using oxygen therapy

2.5.1 Safety Variables

Safety assessments performed during the study included measurement of vital signs and monitoring of adverse events and adverse IP reactions.

3. SUBJECT POPULATIONS

3.1. Population Definitions

The following subject populations will be evaluated and used for presentation and analysis of the data:

- Intent-to-treat (ITT) Population: The ITT Population includes all randomized patients.
- Modified intent-to-treat (mITT): Randomized patients who did not have a primary outcome event within 24 hours of receiving IP.

- Per-protocol (PP) Population: Randomized patients who adhered to more than 80% of the assigned therapy.
- Safety Population: Randomized patients who received at least 1 dose of IP.

The ITT Population is typically the primary population for the analysis of efficacy parameters. A subset of efficacy parameters will be evaluated for the mITT and PP populations (see Section 4.6).

The Safety Population is typically the primary population for the analysis of safety endpoints.

3.2. Protocol Violations

At the discretion of the sponsor, major protocol violations as determined by a review of the data prior to unblinding of the study results and the conduct of statistical analyses may result in the removal of a subject's data from the per-protocol population.

The sponsor, or designee, will be responsible for producing the final protocol violation file (formatted as a Microsoft Excel file), in collaboration with RainCity Analytics and the data monitoring group as applicable; this file will include a description of the protocol violation, and clearly identify whether or not this violation warrants exclusion from the per-protocol population. This file will be finalized prior to hard database lock.

All protocol violations will be presented in the data listings. The major protocol violations will include:

- Failure to obtain informed consent prior to initiation of study-related procedures
- A research subject met withdrawal criteria or wished to withdraw from the study but was not withdrawn
- Inappropriate IP dosage

- Inappropriate randomization
- Inadvertent loss of samples or data
- Other major violation

The minor protocol deviations include:

- Concomitant medication
- Non-compliance to study procedures
- Visit made outside of the visit window

4. STATISTICAL METHODS

4.1. Sample Size Justification

The sample size of 681 patients per arm has been chosen for each experimental group to achieve 80% power with 0.05 two-sided Type 1 error for a pairwise comparison against the control to detect minimum treatment efficacy defined by 37.5% relative risk reduction (RRR) of preventing hospitalization assuming a control event rate (CER) of 15%.

4.1.1. Sample Size Re-Assessment

Sample size re-assessment procedures are described in section 4.3.

4.2. General Statistical Methods and Data Handling

4.2.1. General Methods

All output will be incorporated into Microsoft Excel or Word files, sorted and labeled according to the International Conference on Harmonisation (ICH) recommendations, and formatted to the appropriate page size(s).

Tabulations will be produced for appropriate demographic, baseline, efficacy, and safety parameters. For categorical variables, summary tabulations of the number and percentage within each category (with a category for missing data) of the parameter will be presented. For continuous variables, the mean, median, standard deviation, minimum and maximum values will be presented. Time to event data will be summarized using Bayesian and/or frequentist Kaplan-Meier estimates.

Bayesian bivariate analysis will be performed on the primary efficacy outcome. Bayesian and/or frequentist bivariate analysis will be performed secondary efficacy endpoints outlined in section 2.5. Summary statistics will be presented, as well as their corresponding 95% credible intervals.

4.2.2. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software (Version 9.4) or R (Version 4), unless otherwise noted. Medical History and adverse events will be coding using MedDRA version 23.0. Concomitant medications will be coded using World Health Organization (WHO) Drug dictionary (version March 1, 2020).

4.2.3. Statistical Analysis Details

4.2.3.1. Bayesian inference for dichotomous outcomes with covariate adjustment

Here we may use the generalized linear model framework for binary outcome with the logistic link function, namely

$$y_i \sim \text{Binom}(1, p_i)$$

with

$$\log \frac{p_i}{1-p_i} = x_i^{\mathrm{T}}\beta + \gamma T_i,$$

using the same notations as before. Assigning a noninformative prior distribution $p(\beta, \gamma) \propto 1$, the logarithm of the posterior distribution (after some simple algebra) is given by

$$\log p(\beta, \gamma|\boldsymbol{y}) = \log p(\beta, \gamma) + \sum_{\mathrm{i}=1}^{\mathrm{n}} \log p(y_i|\beta, \gamma)$$

$$= \sum_{i=1}^{n} \{y_i(x_i^T\beta + \gamma T_i) - log[1 + exp(x_i^T\beta + \gamma T_i)]\}, (1)$$

and a random sample from $p(\beta, \gamma|\boldsymbol{y})$ may then be drawn using Markov Chain Monte Carlo (MCMC) such as the Metropolis-Hastings algorithm[1] using (5**Error! Reference source not found.** O r Gibbs sampling using any software for Hierarchical Bayesian modeling such as BUGS[2] or JAGS[3]. Inference on the treatment effect will then follow the same procedure as in the numeric case. The treatment effects on dichotomous outcomes will be estimated using R.

4.2.3.2. Bayesian inference for time-to-event outcomes with covariate adjustment

We assume here that the data satisfies the Cox proportional hazards assumption, that is –

$$\log h(t; x, T, \beta, \gamma) = \log h_0(t) + x^{\mathrm{T}}\beta + \gamma T,$$

where $h_0(t)$ is the baseline hazard function and $h(t; x, T, \beta, \gamma)$ is the hazard function of an individual with covariate vector x who was assigned to treatment T, evaluated at time t.

We model the baseline hazard as a piecewise constant function

$$h_0(t;\boldsymbol{\lambda}) = \sum_{j=1}^{J} \lambda_j \, \mathbb{I}\big(t \in (a_{j-1}, a_j]\big) \quad (1)$$

as in Ibrahim et al.[4] for $\lambda_1, \ldots \lambda_J \geq 0$ and some partition $a_0 = 0 < a_1 < \cdots < a_J = t_{max}$ of the real line with t_{max} the end of the follow-up period. This induces the baseline survival function

$$S_0(t;\boldsymbol{\lambda}) = -\sum_{j=1}^{J} \lambda_j \, (t - a_{j-1}) \mathbb{I}\big(t \in (a_{j-1}, a_j]\big), \quad (2)$$

and, denoting $\boldsymbol{t} = [t_1, \ldots, t_n]^{\mathrm{T}}$ the vector of event times, the log-likelihood function is given by

$$\ell(\boldsymbol{t};\beta,\gamma,\boldsymbol{\lambda}) = \sum_{i=1}^{n} [(1-\delta_i) log\ h_0(t_i;\boldsymbol{\lambda}) + x_i{}^T\beta + \gamma T_i + \log S_0(t_i;\boldsymbol{\lambda}) + \exp(x_i{}^T\beta + \gamma T_i)], \quad (3)$$

where δ_i is an indicator assuming the value of 1 if the i^{th} observation was right-censored and 0 otherwise. We may then assign an improper, independent prior $p(\beta,\gamma,\boldsymbol{\lambda}) \propto \prod_{j=1}^{J} \lambda_j^{-1}$, and proceed to generate an MCMC sample from the posterior distribution, using the log-posterior

$$\log p(\beta,\gamma,\boldsymbol{\lambda}|\boldsymbol{t}) = \text{const} + \ell(\boldsymbol{t};\beta,\gamma,\boldsymbol{\lambda}) - \sum_{j=1}^{J} \log \lambda_j$$

within a Metropolis-Hastings scheme. Inference on treatment efficacy, as always, will be based on the $100(1-\alpha)\%$ credible interval for γ.

The treatment effects on time-to-event outcomes will be estimated using R.

4.2.3.2.1. Checking for proportional hazards assumption for time-to-event analyses

The proportional hazards assumption will first be checked visually inspecting the Kaplan-Meier plots of the survival function versus the survival time. Plots of log(-log(survival)) versus log(survival) will also be generated to check the proportional hazards assumption.

4.2.4. Multiple Comparisons/Multiplicity

All subgroup analyses are considered exploratory for this stage of the TOGETHER trial and thus do not require adjustment for multiplicity. Treatment comparisons in focus are solely experimental versus placebo comparisons and for each treatment a matching placebo exists. Thus, all placebo comparisons can be considered approximately independent and therefore not requiring multiplicity adjustments. Lastly, multiplicity due to repeated testing is handled with Bayesian stopping rules (see section 4.3).

4.2.5. Subpopulations

The following subpopulations will be considered for subgroup analyses:

- Age:
    - ≥50 years or <50 years
- Sex: Male or female
- Time from onset of symptoms:
    - ≥ 120 hours or < 120 hours
- Comorbidity in screening
    - Diabetes mellitus (yes or no);
    - Cardiovascular disease (yes or no);
    - Lung disease (yes or no);
    - Immunosuppressed patients / use of corticosteroid therapy (Yes or No);
    - Other special categories (solid organ transplantation, end-stage kidney disease).

4.2.6. Withdrawals, Dropouts, Loss to Follow-up

Subjects who withdrew from the study will not to be replaced.

4.2.7. Missing, Unused, and Spurious Data

All data recorded on the CRF will be included in data listings that will accompany the clinical study report. Due to the design of the study and retention activities, we expect to be able to measure outcomes on all or the vast majority of participants. Multiple imputation will be employed where statistical models require adjustment for baseline covariates with up to 20% missing values. No multiple imputation of outcomes will be performed.

4.2.8. Visit Windows

Refer to the protocol for treatment periods, follow-up periods, and visit windows, as these vary by IP.

4.3. Interim Analyses

Three interim efficacy analyses are planned. Assuming a uniform prior assigned to the different event rates, a total sample size of 681 patients per arm, a CER of 15%, and a RRR equals to 37.5%, we will do an interim analysis after observing 25%, 50% and 75% of the maximum number of patient outcomes, as well as at the trial completion. The posterior efficacy threshold will be 97.6% and the futility threshold will be 20%, 40% and 60%. Intervention arms(s) showing a posterior probability of efficacy crossing either boundary, will be stopped for either reason. These superiority and futility thresholds were determined based on 200,000 simulation runs in which different values of the RRR were considered (0%, 20%, and 37.5%). A description of this interim

analysis in an event-based Bayesian adaptive trial and accompanying illustrating example can be found in the appendix of this document.

4.3.1. Sample Size Re-Assessment for Brazil

Given that intervention arm(s) are neither superior or futile at the time of the first interim analysis for binary outcome analysis, sample size re-assessment will be performed based on COVID-19 related hospitalization or emergency room visit (for patients under observation for 6 hours or more). For binary outcomes, the sample size and the observed number of events in the control and treatment arms at the time of interim analysis will be used to calculate the future sample size required to achieve 90% BPP. The technical details can be found in Harari and colleagues' paper published in the Pharmaceutical Statistics.[5]

4.3.2. Borrowing Strengths from External Studies

Should individual patient data (IPD) from other relevant studies become available, we may use Empirical Bayes IPD meta-analysis[6] to borrow information from the treatment effects emerging from these studies. This is effectively a random effect Bayesian model that results in simultaneous shrinkage of the treatment effect estimates reported in the various studies toward the meta-analysis estimate, while still providing standalone estimates. Schoenfeld et al. have shown[7] that this approach is, in some ways, equivalent to the power prior approach of Ibrahim and Chen[8], whereby historical studies are assigned a fractional weight whose magnitudes correspond to the consistency of their data with that of the study they are thought to inform. Under the Empirical Bayes IPD meta-analysis model, covariates that may explain differences between studies will be retrieved, converted to similar scales and be included in the model for statistical adjustments. The selection of covariates will be pseudo informal, partially guided by expert advice and partially guided for *forward selection*.

4.4. Subject Disposition

A tabulation of subject disposition will be tabulated, including the number screened, the number dosed with each experimental drug(s) and control, the number in each subject population for analysis, the number that withdrew prior to completing the study, and reasons for withdrawal.

A by-subject listing of study completion information, including the reason for premature study withdrawal, if applicable, will be presented.

4.5. Demographic and Baseline Characteristics

Baseline, demographic and medical history information will be summarized for the *<XX, XX and XX populations>* using descriptive statistics. No formal statistical comparisons will be performed.

Demographic and Baseline data will be provided in data listings.

4.6. Efficacy Evaluation

Efficacy analysis will be conducted using the ITT, mITT, and PP Populations as outlined below. The two primary endpoints will be analyzed as a composite. Secondary endpoints will be assessed within 28 days of randomization, unless stated differently.

Table 1: Objectives and Endpoints

Objectives	Endpoints
Primary	
• Reduction in the need for emergency room visits due to the clinical worsening of COVID-19 and keeping the participant under observation for > 6 hours within 28 days of	• Proportion of those in need of emergency care under observation for more than 6 hours for COVID-19

randomization in acutely affected patients and with evidence of high-risk for complications	
• Reduction in the need for hospitalization due to the progression of COVID-19 (worsening of viral pneumonia) or complications within 28 days of randomization in acutely affected patients and with evidence of high-risk for complications	• Proportion of those in need of hospitalization due to progression of COVID-19 and/or complications
Secondary	
• Reduction in all-cause and COVID-19-related mortality up to 28 days after randomization	• Proportion of participants who died due to any cause • Proportion of participants who died due do COVID-19-related complications
• To test the efficacy of experimental interventions to reduce SARS-CoV-2 viral load at day 3 and day 7	• Proportion of persons with clearance of SARS-CoV-2 from nasal swabs or saliva, defined as 1 negative swab • Proportion of days with SARS-CoV-2 detected from mid-nasal swabs by PCR • Change in viral load on day 3 and day 7 compared to baseline
• To test whether any of the IPs decrease time to resolution for symptomatic SARS-CoV-2 infection / COVID-19 disease	• Number of days with respiratory symptoms • Time to clinical improvement (up to 28 days of randomization), defined as the first day on which the participant reports a score of 0 on the WHO Clinical Worsening Scale • Proportion of patients with change in WHO clinical worsening scale over the treatment period and over the follow-up period • Proportion of patients with change in the Wisconsin Upper Respiratory Symptom Survey (WURSS-21) over the treatment period and over the follow-up period

• To test whether any of the IPs are associated with decreased time to mortality within 28 days of randomization	• Time to death for any cause, including COVID-19
• To test whether any of the IP are associated with decreased hospitalization for any causes	• Hospitalized for any cause, including COVID-19
• To test whether any of the IPs are associated with the time of hospitalization/urgent care due to COVID-19	• Time to hospitalization/urgent care due to COVID-19 progression
• To test the whether any of the IPs are associated with quality-of-life (QoL)	• Change in quality of life measured by PROMIS Global-10 from baseline to Day 14 and Day 28
• To test whether any of the IPS are associated with improved cognitive ability at day 28	• Telephone Interview for Cognitive Status (TICS) questionnaire score at day 28
Exploratory	
• To test whether any of the IPs are associated with oxygen saturation ≤ 93	• If oxygen saturation ≤ 93 over the course of the follow-up period
• To test whether any of the IPS are associated with the number of days spent in the intensive care unit	• Number of days spent in an intensive care unit
• To test whether any of the IPS are associated with the number of days on invasive mechanical ventilation	• Number of days on invasive mechanical ventilation
• To test whether any of the IPS are associated with the number of days in hospital	• Number of days of hospitalization
• To test whether any of the IPS are associated with the number of days on the hospital ward	• Number of days of hospitalization in the ward
• To test whether any of the IPS are associated with the number of days using oxygen therapy	• Number of days using oxygen therapy

4.7. Safety Analyses

Safety analyses will be conducted using the Safety Population.

4.7.1. Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and displayed in tables and listings using System/Organ/Class (SOC) and Preferred Term.

Analyses of adverse events will be performed for those events that are considered treatment emergent, where treatment emergent is defined as any adverse event with onset after the administration of study medication or any event that was present at baseline but worsened in intensity or was subsequently considered drug-related by the investigator through the end of the study.

Adverse events are summarized by subject incidence rates, therefore, in any tabulation, a subject contributes only once to the count for a given adverse event (SOC or preferred term).

The number and percentage of subjects with any treatment-emergent adverse events assessed by the Investigator as related to treatment (definite, probable, or possible relationship), and with any serious adverse event will be summarized by treatment group and overall. In these tabulations, each subject will contribute only once (i.e., the most related occurrence or the most intense occurrence) to each of the incidence rates in the descriptive analysis, regardless of the number of episodes.

No formal hypothesis-testing analysis of adverse events incidence rates will be performed.

All adverse events occurring on study will be listed in subject data listings.

By-subject listings also will be provided for the following: subject deaths; serious adverse events; and adverse events leading to withdrawal.

4.7.2. Temperature and Arterial Oxygen Saturation

Temperature and arterial oxygen saturation will be summarized descriptively, including the number and percent of subjects with normal, abnormal, and clinically significant results at Baseline. All temperature and arterial oxygen saturation data for each subject will be provided in data listings.

4.7.3. Electrocardiogram

ECG results will be summarized descriptively, including the number and percent of subjects with normal, abnormal and clinically significant abnormal results at Baseline. All ECG data for each subject will be provided in data listings.

4.7.4. Concomitant Medications

Concomitant medications will be coded using the WHO Drug dictionary. Results will be tabulated by Anatomic Therapeutic Class (ATC) and preferred term. The use of concomitant medications will be included in by-subject data listing.

4.7.5. Vaccination Status

Vaccination status will be tabulated and will be included in by-subject data listing. It may also be used in a sensitivity analysis.

4.8. CHANGES TO PLANNED ANALYSES

As of this date, there have been no changes between the protocol-defined statistical analyses and those presented in this statistical plan. Post-protocol analyses may be added later due to the rapid evolvement of discoveries for COVID-19, but these will be considered exploratory.

REFERENCES

1. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis: CRC press; 2013.
2. Sturtz S, Ligges U, Gelman AE. R2WinBUGS: a package for running WinBUGS from R. 2005.
3. Plummer M. JAGS: Just another Gibbs sampler. 2004.
4. Ibrahim JG, Chen MH, Sinha D. B ayesian Survival Analysis. Wiley StatsRef: Statistics Reference Online. 2014.
5. Harari O, Hsu G, Dron L, Park JJH, Thorlund K, Mills EJ. Utilizing Bayesian predictive power in clinical trial design. Pharm Stat. 2020.
6. Raudenbush SW, Bryk AS. Empirical Bayes Meta-Analysis. Journal of Educational Statistics. 1985;10(2).
7. Schoenfeld DA, Hui Z, Finkelstein DM. Bayesian design using adult data to augment pediatric trials. Clin Trials. 2009;6(4):297-304.
8. Chen M-H, Ibrahim JG. Power prior distributions for regression models. Statistical Science. 2000;15(1):46-60.

APPENDIX

1. Interim analysis in an even-based Bayesian adaptive clinical trial

Suppose that we have I active treatments, such that

$$y_{ij}|p_i \sim \text{Binom}(n_{ij}, p_i), \qquad i = 1, \dots, I,$$

where $j = 1, \dots, J$ denotes the number of interim analysis, n_{ij} is the number of patients randomized to the i^{th} arm, and p_i is the corresponding event rate. If we then assign a prior distribution

$$p_i \sim \text{Beta}(a, b),$$

Then by conjugation we obtain

$$p_i|y_{ij} \sim \text{Beta}(a + y_{ij}, b + n_{ij} - y_{ij}). \quad (1)$$

Control comparisons will then be based on *posterior efficacy*, namely

$$\Pr(p_i < p_{\text{ctrl}}|y_{ij}, y_{\text{ctrl}}),$$

which can be handily calculated by drawing independent Monte Carlo samples from the posterior distributions of the two arms using (1). Posterior inference on the relative risk reduction (RRR) can then be derived using the relationship

$$R_i = 1 - \frac{p_i}{p_{\text{ctrl}}}. \quad (2)$$

Proposed design for a 4-arm trial

- Perform interim analysis when 171, 342, and 513 patient outcomes have been recorded for each of the trial arms – corresponding to 25%, 50% and 75% of the maximum enrollment - and a final analysis when all 681 patient outcomes have been registered.
- Use $a = b = 1$ for beta prior distribution of all arms, corresponding to a uniform distribution.

- Stop early for efficacy if the posterior probability of efficacy exceeds 97.6%. The same test is conducted at the end of the trial if no early stopping rules are triggered.
- Stop early for futility if the posterior probability of efficacy is smaller than 20% in the first analysis; 40% in the second analysis; 60% in the third analysis.

The operating characteristics of this design, as determined by 200,000 simulation runs, are as follows.

Table *2*: Operating Characteristics for the Proposed Bayesian Design

Arm	RRR	Power	Mean % of patients (SD)	Pr(Stop at 1)	Pr(Stop at <=2)	Pr(Stop at <=3)
1	37.50%	84.80%	64.1 (29.6)	25.20%	53.50%	73.80%
2	20%	30.70%	84.2 (26.0)	10.40%	23.90%	39.20%
3	0%	2.50%	71.1 (28.4)	16.80%	42.50%	67.00%

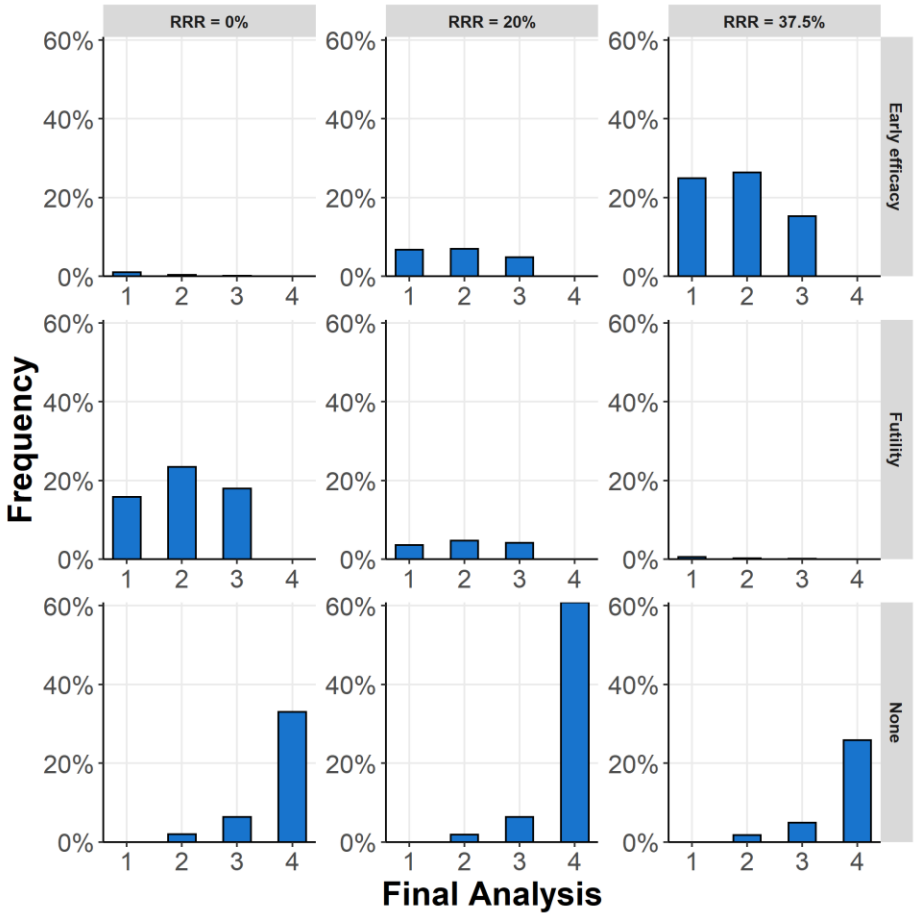
From Table 2 it is evident that the type I error rate of this trial design is 2.5% (one-sided). Note that a treatment with a relative risk reduction of 37.5% will – on average – require only two-thirds of the maximum sample size of 681.

The simulation run also allows us to evaluate the frequency of the reasons for early stopping under each scenario. The details are given in Table 3 and Figure 2.

Table *3*: Distribution of the Causes for Early Stopping Under Various Scenarios

RRR	Stopping Reason	Analysis 1	Analysis 2	Analysis 3	Analysis 4
0%	Early efficacy	1.00%	0.30%	0.20%	0.00%
	Futility	15.80%	23.50%	18.00%	0.00%
	None	0.00%	1.90%	6.40%	32.90%
20%	Early efficacy	6.70%	6.90%	4.90%	0.00%
	Futility	3.60%	4.70%	4.10%	0.00%
	None	0.00%	1.80%	6.30%	60.90%
37.5%	Early efficacy	24.70%	26.40%	15.30%	0.00%
	Futility	0.50%	0.20%	0.20%	0.00%
	None	0.00%	1.70%	4.90%	26.20%

Figure 2: a graphical illustration of Table 2.



Example: a simulated 4-arm trial

To illustrate the application of the decision rules in the proposed design, we drew one instance of the data according to the 15% control event rate and the specified effect sizes. The resultant trial consisted of 4 analyses: three interim analyses and one final analysis. Table 4 monitors the cumulative number of events over time for each arm. Note that already at the first interim look, arm 3 was dropped due to hitting a futility rule after 171 patient outcomes were observed. Two

analyses and 342 patients later, arm 1 crossed the efficacy threshold, following which no more patients were randomized to it. Arm 2 continued until all 681 patient outcomes were recorded and did not hit either decision boundary. This is illustrated in Figure 3.

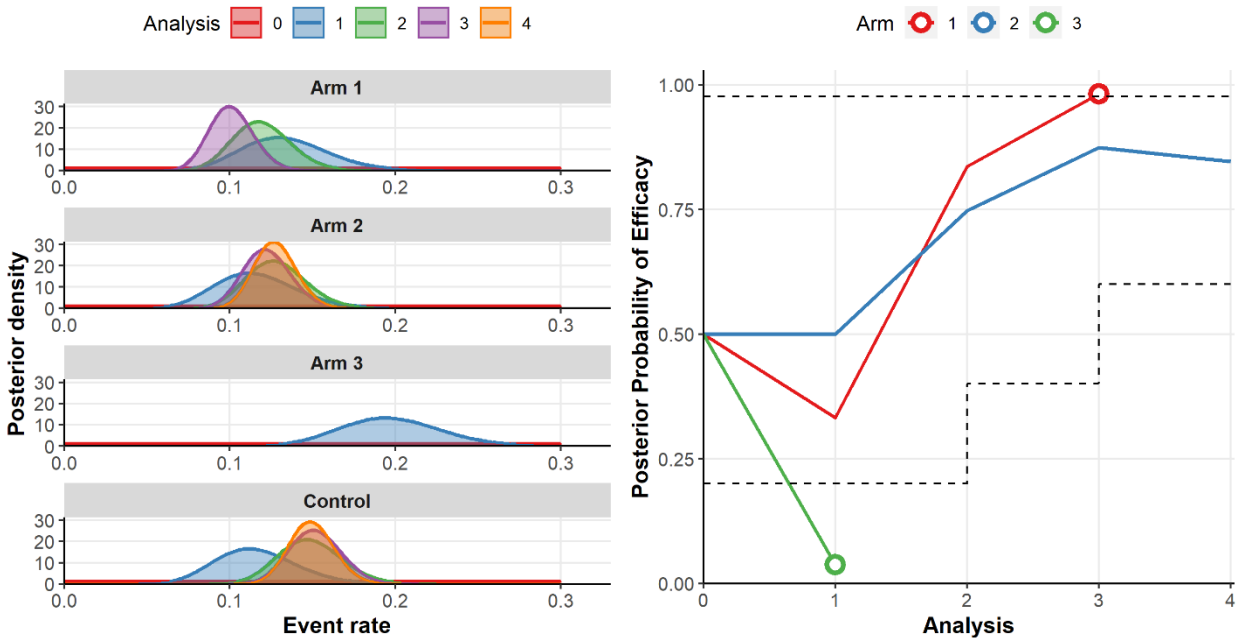
An alternative form of Bayesian data monitoring is an examination of the posterior densities of the different p_i parameters ($i \in \{\text{Control}, 1,2,3\}$) over time as in Figure 3, starting from a uniform prior reflecting ignorance, and with centers that change over time and with reduced uncertainty, as more evidence is gathered.

Table 4 also records the proportion of patient outcomes out of the maximum number of 681 that were recorded in the trial, as well as the final decision (efficacious/non-efficacious) and a 95% interval estimate of the relative risk reduction, based on the relationship (2).

Table *4*: Data and Bayesian Analysis of a Single Simulated Trial

Arm	Analysis 1 events	Analysis 2 events	Analysis 3 events	Analysis 4 events	Stopping Reason	% patients	Pr(Efficacy|Data)	Efficacy	RRR
1	15	28	45		Early efficacy	75.3	98.60%	Yes	36.3% [9.3%; 55.9%]
2	21	44	67	90	Enrolled all	100	52.60%	No	1.0% [-30.6%; 24.2%]
3	28	56			Futility	50.2	15.90%	No	-23.5% [-78.2%; 11.7%]
Control	22	45	71	91		100			

*Figure 3: the evolution of the posterior density of the event rate of each arm by interim analysis for the simulated example and posterior efficacy of all active arms by interim analysis**



**The dashed lines mark the decision boundaries.*