Statistical Analysis Plan

INCB 18424-369/ NCT04377620

A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study to Assess the Efficacy and Safety of Ruxolitinib in Participants With COVID-19–Associated ARDS Who Require Mechanical Ventilation (RUXCOVID-DEVENT)

IND Number:	**149,270**
Sponsor:	**Incyte Corporation** **1801 Augustine Cut-Off** **Wilmington, DE 19803** **United States**
Protocol Version:	**Amendment 1 dated 02 JUN 2020**
CRF Approval Date:	**30 SEP 2020**
SAP Version:	**Amendment 3**
SAP Author:	[redacted]
Date of Plan:	**22 FEB 2021**

This study is being conducted in compliance with Good Clinical Practice, including the archiving of essential documents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Term
AE	adverse event
ALT	alanine aminotransferase
ARDS	acute respiratory distress syndrome
AST	aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical Classification System
BID	twice daily
BMI	body mass index
COVID-19	coronavirus disease 2019
CRF	case report form
CRP	C-reactive protein
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
ECMO	extracorporeal membrane oxygenation
eCRF	electronic case report form
FiO_2	percentage of inspired oxygen
ICU	intensive care unit
ITT	intent to treat
LDA	longitudinal data analysis
MedDRA	Medical Dictionary for Regulatory Activities
NCI	National Cancer Institute
PaO_2	partial pressure of arterial oxygen
PLB	placebo
PP	per protocol
PT	preferred term
SAP	Statistical Analysis Plan
SARS-CoV-2	severe acute respiratory syndrome coronavirus 2
SOC	system organ class
SoC	standard of care
SOFA	sequential organ failure assessment
TEAE	treatment-emergent adverse event
ULN	upper limit of normal
WHO	World Health Organization

1. INTRODUCTION

This is a randomized, double-blind, placebo-controlled, 29-day, multicenter study to assess the efficacy and safety of ruxolitinib 5 mg BID + SoC therapy and 15 mg BID + SoC therapy compared with placebo + SoC therapy in participants aged ≥ 12 years with COVID-19–associated ARDS who require mechanical ventilation. Participants will be assigned in a 2:2:1 ratio to receive ruxolitinib 5 mg BID, 15 mg BID, or matching-image placebo for an initial period of 14 days; participants randomized to placebo will be randomized 1:1 to either receive matching placebo for 5 mg BID (1 tablet BID) or matching placebo for 15 mg BID (3 tablets BID). On Day 15, if in the opinion of the investigator the benefit/risk is appropriate, then continued treatment up to 28 days is permitted with medical monitor approval. Study treatment will be given in combination with SoC therapy according to the investigator's clinical judgment. Study treatment will be administered via an enteric feeding tube while the participant is intubated. Randomization will be stratified by ARDS severity (severe [$PaO_2/FiO_2 \leq 100$] vs mild/moderate [$PaO_2/FiO_2 > 100$ and ≤ 300]) and investigative site. Sections 2 and 5 of the Protocol provide a detailed description of the investigational product, target patient population, rationale for doses to be examined, and potential risks and benefits of treatment with ruxolitinib/placebo.

The purpose of this SAP is to provide details of the statistical analyses that have been outlined in the INCB 18424-369 Protocol.

The details of the analysis methodology of biomarkers and pharmacodynamics and results will appear in a separate report.

2. STUDY INFORMATION, OBJECTIVES, AND ENDPOINTS

2.1. Protocol and Case Report Form Version

This SAP is based on INCB 18424-369 Protocol Amendment 1 dated 02 JUN 2020 and CRFs approved on 30 SEP 2020. Unless superseded by an amendment, this SAP will be effective for all subsequent Protocol amendments and eCRF versions.

2.2. Study Objectives and Endpoints

Table 1 presents the objectives and endpoints.

Table 1: Objectives and Endpoints

Objectives	Endpoints
Primary	
To evaluate the 28-day mortality rate of ruxolitinib 5 mg BID + SoC therapy and ruxolitinib 15 mg BID + SoC therapy compared with placebo + SoC therapy in participants with COVID-19–associated ARDS who require mechanical ventilation.	Proportion of participants who have died due to any cause through Study Day 29.
Secondary	
To evaluate the efficacy of ruxolitinib 5 mg BID + SoC therapy and ruxolitinib 15 mg BID + SoC therapy compared with placebo + SoC therapy on in-hospital outcomes in participants with COVID-19–associated ARDS who require mechanical ventilation.	At Day 29, the number of ventilator-free days, ICU-free days, oxygen-free days, vasopressor-free days, and hospital-free days will be summarized by treatment group.
To evaluate the efficacy of ruxolitinib 5 mg BID + SoC therapy and ruxolitinib 15 mg BID + SoC therapy compared with placebo + SoC therapy using a 9-point ordinal scale at Study Days 15 and 29 in participants with COVID-19–associated ARDS who require mechanical ventilation.	Ordinal scale: • Percentage of participants with at least 2-point improvement in clinical status at Day 15 and at Day 29. • Percentage of participants with at least 1-point improvement in clinical status at Day 15 and at Day 29. • Time to improvement from baseline category to earliest 1-point improvement in the ordinal scale. • Percentage of participants in each 9-point ordinal scale category at Day 29. • Mean change in the 9-point ordinal scale from baseline to Days 15 and 29.

Table 1: Objectives and Endpoints (Continued)

Objectives	Endpoints
To evaluate the efficacy of ruxolitinib 5 mg BID + SoC therapy and ruxolitinib 15 mg BID + SoC therapy compared with placebo + SoC therapy on change from baseline SOFA score in participants with COVID-19–associated ARDS who require mechanical ventilation.	Change from baseline to Day 3, 5, 8, 11, 15, and 29 in SOFA score.
To evaluate the safety of ruxolitinib 5 mg BID + SoC therapy and ruxolitinib 15 mg BID + SoC therapy compared with placebo + SoC therapy in the treatment of participants with COVID-19–associated ARDS who require mechanical ventilation.	Number and proportion of participants with treatment-related side effects (as assessed by CTCAE v5.0) and SAEs; includes clinically significant changes in laboratory measures and vital signs.
Exploratory	
To evaluate the efficacy of ruxolitinib + SoC therapy compared with placebo + SoC therapy in proportion of participants requiring IL-6, IL-1, GM-CSF, or JAK-directed therapies.	Proportion of participants requiring treatments with IL-6, IL-6RA, IL-1RA, IL-1 beta, GM-CSF, BTK, or JAK directed therapies by Day 15 and by Day 29.
To evaluate ruxolitinib + SoC therapy compared with placebo + SoC therapy on change in inflammatory markers (when available).	Change from baseline to Day 15 and to Day 29 in the following clinical chemistry measurements: • Serum ferritin • CRP • D-dimer • Procalcitonin • IL-6
To evaluate ruxolitinib + SoC therapy compared with placebo + SoC therapy on change in COVID-19 virologic and serologic parameters.	Change from baseline to Day 15 and Day 29 in viral load and anti–SARS-CoV-2-antibody titer.

3. STUDY DESIGN

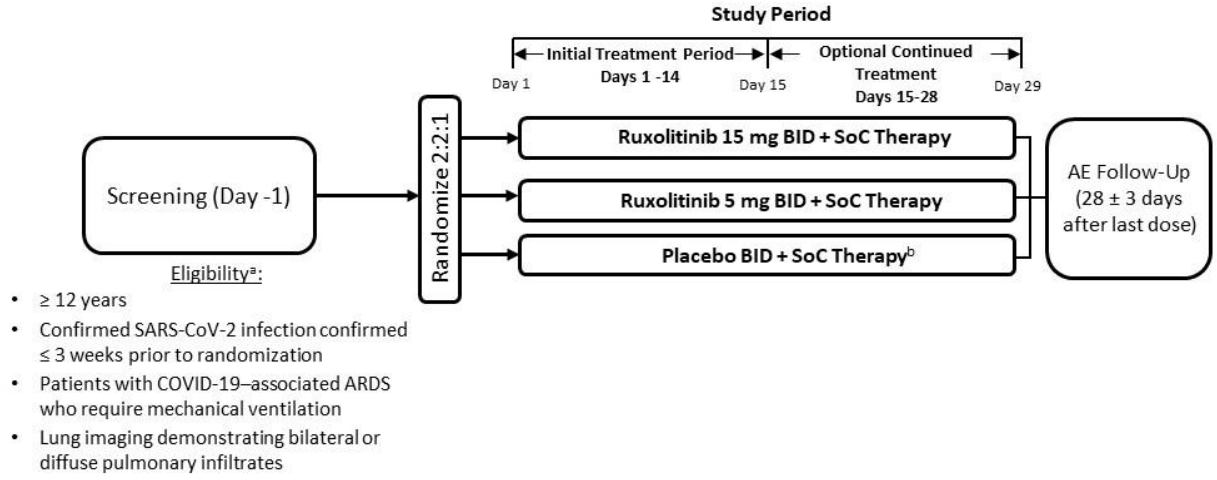
This is a randomized, double-blind, placebo-controlled, 29-day, multicenter study to assess the efficacy and safety of ruxolitinib 5 mg BID + SoC therapy and 15 mg BID + SoC therapy compared with placebo + SoC therapy in participants aged ≥ 12 years with COVID-19–associated ARDS who require mechanical ventilation (see Figure 1).

Participants aged ≥ 12 years with SARS-CoV-2 infection confirmed ≤ 3 weeks before randomization by any test with local regulatory approval, hospitalized with COVID-19–associated ARDS (demonstrated by chest x-ray or chest computerized tomography and $PaO_2/FiO_2 \leq 300$ mmHg) who have been intubated and are receiving mechanical ventilation. Participants who are unlikely to survive for > 24 hours after randomization (in the opinion of the treating investigator); who have active tuberculosis or active and uncontrolled bacterial, fungal, viral, or other infection (besides SARS-CoV-2); or who are currently receiving ECMO support will be excluded from the study.

The study will include the following:

- Screening period of 1 day.
- Study period of 29 days. Treatment with ruxolitinib 5 mg BID/placebo or ruxolitinib 15 mg BID/placebo will initially be administered for 14 days. If in the opinion of the treating investigator the benefit/risk is appropriate for the participant, then continued treatment up to 28 days permitted with approval from the sponsor medical monitor. Approval must be requested ≥ 24 hours before initiation of dosing on Day 15.
- AE follow-up assessment 28 (± 3) days following the last dose of study treatment.

Figure 1: Study Design



[a] Not inclusive; see Protocol Section 6 for complete eligibility criteria.
[b] Placebo arm will be randomized between a 5 mg BID (1 tablet/dose) and 15 mg BID (3 tablets/dose).

Eligible participants will be randomized after completing the screening period (Day -1) and having been confirmed eligible. Screening and randomization may occur on the same day. Participants will be assigned in a 2:2:1 ratio to receive ruxolitinib 5 mg BID, 15 mg BID, or matching-image placebo for an initial period of 14 days (refer to Protocol Section 7.1);

participants randomized to placebo will be randomized 1:1 to receive either matching placebo for 5 mg BID (1 tablet BID) or matching placebo for 15 mg BID (3 tablets BID). On Day 15, if in the opinion of the investigator the benefit/risk is appropriate, then continued treatment up to 28 days permitted with medical monitor approval. There is no crossover between treatment groups. Study treatment will be given in combination with SoC therapy according to the investigator's clinical judgment. Study treatment will be administered via an enteric feeding tube (refer to Protocol Section 7.1) while the participant is intubated. Randomization will be stratified by ARDS severity (severe [$PaO_2/FiO_2 \leq 100$] vs mild/moderate [$PaO_2/FiO_2 > 100$ and ≤ 300]) and investigative site.

3.1. Randomization

Participants will be randomized 2:2:1 to receive ruxolitinib 5 mg BID, 15 mg BID, or matching-image placebo for an initial period of 14 days. Placebo participants will be randomized 1:1 to receive either the 5 mg BID image or the 15 mg BID image. Block randomization with stratification for ARDS severity (severe [$PaO_2/FiO_2 \leq 100$] vs mild/moderate [$PaO_2/FiO_2 > 100$ and ≤ 300]) and investigational site will be used.

3.2. Control of Type I Error

The Type 1 error level for the primary endpoint is 1-sided 2.5%. A fixed sequence testing procedure will be used to test the 15 mg BID group against placebo first, and if that is significant, then the 5 mg BID group will be tested against placebo. If the test for the 15 mg BID group is not significant, the 5 mg BID group will not be tested.

If the null hypotheses are rejected for both ruxolitinib treatment groups when compared with placebo for the primary endpoint, then a statistical comparison will be performed between the 2 ruxolitinib groups.

No multiplicity adjustments will be applied to other endpoints. Secondary efficacy variables will be tested at a 0.05 level and will use a 2-sided test. All exploratory variables will use a 2-sided, 5% alpha with no adjustment for multiplicity.

3.3. Sample Size Considerations

The primary endpoint of this study is the proportion of participants who die due to any cause through Study Day 29. The primary endpoint will be tested for each ruxolitinib treatment group using a logistic regression model including treatment (ruxolitinib vs placebo) and ARDS severity (severe [$PaO_2/FiO_2 \leq 100$] vs mild/moderate [$PaO_2/FiO_2 > 100$ and ≤ 300] as fixed covariates and investigational site as a random (intercept) effect. For placebo randomized 2:2:1 (ruxolitinib 5 mg BID:ruxolitinib 15 mg BID:placebo) assuming a 40% mortality rate for ruxolitinib versus 60% for placebo, a sample size for a pairwise comparison of 100 placebo and 200 ruxolitinib participants (with 500 participants total) will achieve approximately 83% power to detect a statistically significant difference with a nominal 1-sided Type I error of 1.436968%.

Estimated power and Type I error for various scenarios, based on a simulation study of 1000 study replications, are provided in Table 2.

Table 2: Simulation Estimates for Power and Type 1 Error for Various Mortality Rates

Mortality Rates			Probability to Reject H_0 for		
PLB	**5 mg BID**	**15 mg BID**	**5 mg BID**	**15 mg BID**	**Either RUX**
60%	60%	60%	0.015	0.016	0.027
60%	60%	40%	0.012	0.873	0.873
60%	50%	40%	0.268	0.865	0.866
60%	40%	40%	0.868	0.872	0.940
60%	40%	30%	0.861	0.998	0.998

Note: For the simulation, investigational site was drawn from an integer uniform distribution (1, 2, …, 10) and probability (ARDS severe) = 0.5, fitting a mixed logistic regression model with no direct effects for the ARDS severity or site.

Enrollment was halted and SAP Amendment 3 removed the interim analysis. The final analysis will be performed using all randomized participants enrolled at the time of SAP Amendment 3, which is 211 participants.

3.4. Schedule of Assessments

Refer to Protocol Amendment 1 dated 02 JUN 2020 for a full description of all study procedures and assessment schedules for this study.

4. DATA HANDLING DEFINITIONS AND CONVENTIONS

4.1. Scheduled Study Evaluations and Study Periods

4.1.1. Day 1

Day 1 is the date that the first dose of study drug (ruxolitinib/placebo) is administered to the participants.

For randomized participants not treated with any study drug, Day 1 is defined as the date of randomization.

4.1.2. Study Day

If a visit/reporting date is on or after Day 1, then the study day at the visit/reporting date will be calculated as

Day # = (visit/reporting date – Day 1 date + 1).

If the visit/reporting date is before Day 1, then the study day at the visit/reporting date will be calculated as

Day # = (visit/reporting date – Day 1 date).

A study day of –1 indicates 1 day before Day 1.

4.1.3. Baseline Value

Baseline is the last nonmissing measurement obtained before the first administration of ruxolitinib/placebo, unless otherwise defined.

For randomized participants not treated with any study drug, baseline is defined as the last nonmissing assessment before randomization for all parameters.

When scheduled assessments and unscheduled assessments occur on the same day and the time of the assessment or time of first dose is not available, use the following convention to determine baseline:

- If both a scheduled and an unscheduled visit are available on the day of the first dose and the time is missing, use the scheduled assessment as baseline.
- If all scheduled assessments are missing on the day of the first dose and an unscheduled assessment is available, use the unscheduled assessment as baseline.

4.1.4. Last Available Value

The last available value is the last nonmissing measurement obtained after starting ruxolitinib/placebo and within 31 days after the last dose of ruxolitinib/placebo.

4.1.5. Handling of Missing and Incomplete Dates and Times

In general, values for missing dates will not be handled unless methods for handling missing dates are specified in this section or relevant sections. The original reported dates collected on the eCRF should be used in all relevant listings. The following rules will be used for handling partial dates for analyses requiring dates.

When calculating the time since diagnosis, a partial date will be treated as missing and a time will not be calculated. When calculating the time since intubation, a partial date and/or 24-hour time will be treated as missing and a time will not be calculated.

When the date of the last dose is used in deriving variables such as duration of treatment or TEAE flag, a missing or partial date of the last dose will be handled as follows:

- If only the day is missing, then the earlier date of the last day of the month or the date that the participant discontinued treatment will be used.
- Otherwise, the date that the participant discontinued treatment will be used as the date of the last dose.

For relevant efficacy endpoints, a partial date of the death date will be handled as follows in the calculation:

- If mmyyyy for the last known alive date = mmyyyy for the death date, then the death date will be set to the day after the last known alive date.
- If mmyyyy for the last known alive date < mmyyyy for the death date, then the death date will be set to the first day of the death month.
- Otherwise, the partial death date will not be imputed.

4.2. Variable Definitions

4.2.1. Body Mass Index

Body mass index will be calculated as follows:

Body mass index (kg/m^2) = [weight (kg)] / [height (m)]^2.

4.2.2. Prior and Concomitant Medication

Prior medication is defined as any nonstudy medication started before the first dose of ruxolitinib/placebo.

Concomitant medication is defined as any nonstudy medication that is started accordingly:

- Before the date of first administration of ruxolitinib/placebo and is ongoing throughout the study or ends on/after the date of first study drug administration.
- On/after the date of first administration of ruxolitinib/placebo and is ongoing or ends during the course of study.

A prior medication could also be classified as "both prior and concomitant medication" if the end date is on or after the first dose of ruxolitinib/placebo. In the listing, it will be indicated whether a medication is only prior, only concomitant, or both prior and concomitant.

For the purposes of analysis, all medications will be considered concomitant medications unless the medications can unequivocally be defined as not concomitant.

4.2.3. Mean Arterial Pressure

The mean arterial pressure will be calculated as follows:

Mean arterial pressure (mmHg) = [systolic blood pressure (mmHg) + 2 × diastolic blood pressure (mmHg)] / 3.

5. STATISTICAL METHODOLOGY

5.1. General Methodology

Unless otherwise noted, SAS® software (SAS Institute Inc, Cary, NC; v9 or later) will be used for the generation of all tables, graphs, and statistical analyses. Descriptive summaries for continuous variables will include but not be limited to the number of observations, mean, standard deviation, median, minimum, and maximum. Descriptive summaries for categorical variables will include the number and percentage of participants in each category.

5.2. Treatment Groups

This is a randomized, double-blind, parallel treatment group design. Participants will be summarized by treatment groups.

5.3. Analysis Populations

5.3.1. All-Screened Population

The all-screened population will include all participants who provided informed consent (written or verbal).

5.3.2. Intent-to-Treat Population

All participants who are randomized will constitute the ITT population. Treatment groups for this population will be defined according to the treatment assignment at the time of randomization regardless of the actual study drug the participant might take during his/her participation in the study.

The ITT population will be used for the summary of demographics, baseline characteristics, participant disposition, and analyses of all efficacy data.

5.3.3. Per-Protocol Population

The PP population comprises the subset of subjects in the ITT population who are compliant with requirements of the clinical study Protocol, including the following:

- Continued to meet all eligibility criteria on Study Day 1.
- Initiated correct randomized therapy within the Protocol-specified window.
- Did not significantly deviate from the Protocol in terms of missed study assessments, noncompliance with randomized therapy, or prohibited concomitant medications.

The PP population will be used in a supportive sensitivity analysis for the primary efficacy endpoint.

5.3.4. Safety Population

The safety population will include all participants who received at least 1 dose of ruxolitinib/placebo. Treatment groups for this population will be determined according to the actual treatment the participant received regardless of assigned study drug treatment.

All safety analyses will be conducted using the safety population.

6. BASELINE, EXPOSURE, AND DISPOSITION

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

6.1. Demographics, Baseline Characteristics, and Disease History

6.1.1. Demographics and Baseline Characteristics

The following demographics and baseline characteristics will be summarized for the ITT population: age, gender, race, ethnicity, weight, height, and BMI.

6.1.2. Baseline Disease Characteristics and Disease History

The following baseline disease characteristics will be summarized for the ITT population: predefined coexisting disorders, baseline SOFA score, baseline COVID-19 9-point ordinal scale clinical status, baseline ARDS severity (severe [$PaO_2/FiO_2 \leq 100$] vs mild/moderate [$PaO_2/FiO_2 > 100$ and ≤ 300 mmHg]), ARDS severity at randomization, time (in hours) from start of mechanical ventilation to randomization, time (in days) from initial diagnosis to randomization, time (in days) from hospital admission to randomization, baseline CRP (≤ 100 mg/L or > 100 mg/L), baseline D-dimer (≤ 1000 ng/mL or > 1000 ng/mL), baseline ferritin (≤ 300 ng/mL or > 300 ng/mL), BMI at baseline (≥ 30 or < 30), prior or concomitant remdesivir, prior or concomitant convalescent plasma, prior or concomitant corticosteroid, concomitant anti-coagulant, prior biologic, prior or concomitant neutralizing antibodies, and prior or concomitant corticosteroid with prior or concomitant remdesivir.

Time from initial diagnosis will be calculated as follows:

Time since diagnosis (days) = (date of randomization – date of diagnosis + 1).

Time from start of mechanical ventilation will be calculated as follows:

Time from start of mechanical ventilation (hours) = (date and time of randomization – date and time of intubation).

Time from hospital admission to randomization will be calculated as follows:

Time from hospital admission (days) = (date of randomization – date of hospital admission + 1).

6.1.3. Medical History

For participants in the ITT population, medical history will be summarized by assigned treatment group. This summation will include the number and percentage of participants with significant medical history for each body system/organ class as documented on the eCRF.

6.2. Disposition of Participant

The number and percentage of participants, who were treated, who were ongoing with study treatment, who completed study treatment, who discontinued study treatment with a primary reason for discontinuation, who were still in the study, who completed the study, and who withdrew from the study with a primary reason for withdrawal will be summarized for the ITT population. The number of participants randomized by country and/or site will also be provided by treatment group.

6.3. Protocol Deviations

Protocol deviations recorded on the eCRF will be summarized and listed.

6.4. Exposure

For participants in the safety population, exposure to ruxolitinib/placebo will be summarized descriptively as the following:

- **Duration of treatment with ruxolitinib/placebo (days):** date of last dose of ruxolitinib/placebo – date of first dose of ruxolitinib/placebo + 1.
- **Average daily dose of ruxolitinib/placebo (mg/day):** total actual ruxolitinib/placebo dose taken (mg) / duration of treatment with ruxolitinib/placebo (days).
- **Ruxolitinib/placebo dose modifications:** number of participants who had ruxolitinib/placebo dose reduction, escalation, and interruption.

In addition, total daily dose of ruxolitinib/placebo dosing will be summarized by day.

6.5. Study Drug Compliance

As ruxolitinib/placebo dosing will be captured on the appropriate exposure eCRF, overall compliance (%) will be calculated for all participants as follows:

Compliance (%) = 100 × [total dose actually taken] / [total prescribed dose].

The total prescribed dose is defined as the sum of the doses prescribed by the investigator accounting for dose modifications.

6.6. Prior and Concomitant Medication

Prior medications and concomitant medications will be coded using the WHO Drug Dictionary. The number and percentage of participants in the ITT for each prior and concomitant medication will be summarized by WHO drug class and WHO drug PT.

7. EFFICACY

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

7.1. General Considerations

Unless otherwise stated, the strata identified in the randomization process will be used in all efficacy analyses. For statistical analyses of the primary and secondary endpoints, separate models will be fit for the ruxolitinib 5 mg BID versus placebo and ruxolitinib 15 mg BID versus placebo comparisons.

7.2. Efficacy Hypotheses

The primary hypothesis is that ruxolitinib in combination with SoC will reduce 28-day mortality, as measured by odds ratio, compared with placebo in combination with SoC in participants with COVID-19–associated ARDS who require mechanical ventilation. The hypothesis of the study are as follows:

- H_0 (null hypothesis): $\theta_{RUX05} = \theta_{RUX15} = \theta_{PLB}$
- H_A(alternative hypothesis): $\theta_{RUX15} < \theta_{PLB}$ or ($\theta_{RUX05} < \theta_{PLB}$ and $\theta_{RUX15} < \theta_{PLB}$)

The 5 mg BID hypothesis will only be tested if the 15 mg BID hypothesis is significant. In the event that both individual null hypotheses are rejected, then a test for a difference between the 5 mg BID and 15 mg BID treatment groups will be performed for the 28-day mortality endpoint. The hypothesis for this test is:

- H_0 (null hypothesis): $\theta_{RUX05} = \theta_{RUX15}$
- H_A(alternative hypothesis): $\theta_{RUX15} \neq \theta_{RUX05}$

7.3. Analysis of the Primary Efficacy Parameter

7.3.1. Primary Efficacy Analysis

The primary endpoint, 28-day mortality, will be tested using a logistic regression mixed model including treatment group (specific ruxolitinib treatment group vs placebo) and ARDS severity (severe vs mild/moderate) as fixed covariates and investigational site as a random (intercept) effect. The estimation technique for the model fit will use the residual pseudo-likelihood with a subject-specific expansion (SAS Institute 2013) approach and will be implemented in the SAS GLIMMIX procedure. If the model is not estimable due to convergence issues, a fixed effects logistic regression will be performed instead, excluding investigational site from the model. The Wald test for the treatment regression coefficient will test the primary endpoint.

An estimate and 95% CI for the difference in proportions for 28-day mortality for the 2 treatment groups will be provided. The estimate will be based on the logistic regression model (Ge et al 2011) using an average of the difference between the 2 levels of the fixed factor (ARDS severity) weighted for the proportion of the study population within each of the 2 levels

of the fixed factor. The delta method will be used to derive the variance estimate, based on the functional form of the estimate $\gamma^T g(b)$ using $\gamma = (\gamma_1, \gamma_2)$ where

- γ_1 is the proportion of participants in the $b_2 = 0$
- γ_2 is the proportion of participants in the $b_2 = 1$

$\underline{b}$ is the vector of the logistic regression coefficients, and

$$g(\underline{b}) = \begin{matrix} 1/1 + \exp(-b_0 - b_1) - 1/1 + \exp(-b_0) \\ 1/1 + \exp(-b_0 - b_1 - b_2) - 1/1 + \exp(-b_0 - b_2) \end{matrix} .$$

Point estimates and 95% CIs will also be provided for the model-based mortality rates using the delta method.

Participants who are lost to follow-up before Day 29 will not be evaluable for this endpoint. As the participants begin the study intubated, those who remain hospitalized are unlikely to be lost to follow-up, but reasonable efforts will be made to track down missing participants.

7.3.2. Subgroup Analyses for Primary Endpoint

Subgroups will be formed based on the following participant characteristics and baseline variables for those participants whose data are available:

- ARDS severity: severe ($PaO_2/FiO_2 \leq 100$ mmHg) versus mild/moderate ($PaO_2/FiO_2 > 100$ and ≤ 300 mmHg)
- Gender: male or female
- Age group: < 18 years, 18-64 years, or ≥ 65 years
- Baseline CRP: ≤ 100 mg/L or > 100 mg/L
- D-dimer at baseline: ≤ 1000 ng/mL or > 1000 ng/mL
- Ferritin at baseline: ≤ 300 ng/mL or > 300 ng/mL
- Race: White/Caucasian, Black/African-American, Asian, American-Indian/Alaska Native, Native Hawaiian/Pacific Islander, or Other
- Ethnicity: Hispanic/Latino, Not Hispanic/Latino, or Other
- Hours of mechanical ventilation at baseline: ≤ 48 hours or > 48 hours
- BMI at baseline: ≥ 30 or < 30
- Prior or concomitant remdesivir: yes or no
- Prior or concomitant convalescent plasma: yes or no
- Concomitant anticoagulant use: yes or no
- Concomitant corticosteroid use: yes or no
- Prior biologic use: yes or no

7.3.3. Sensitivity and Supportive Analyses for Primary Endpoint

A sensitivity analysis will be conducted in the ITT population, in which nonevaluable participants will be treated as deaths. The primary endpoint will be analyzed using the PP population as a sensitivity analysis to the ITT analysis. In addition, a tipping point analysis will be performed to measure the robustness of the treatment effect under assumptions for participant dropout. The tipping point analysis will use an unstratified 2-sample test of proportion, modifying dropouts within the compared ruxolitinib treatment group as having died and placebo participants as having lived to determine if and at what point in this imputation a statistically significant treatment effect under the base assumptions will lose statistical significance.

Analyses related to the primary endpoints are delineated in Table 3.

Table 3: Primary and Sensitivity Analyses of Primary Endpoints

Endpoint	Statistical Method	Analysis Population	Primary (P) or Supportive (S)
Proportion of participants who have died due to any cause through Study Day 29.	Logistic regression with treatment group and ARDS severity (severe vs mild/moderate) as fixed covariates, and investigational site as a random effect. Wald test for treatment group regression coefficient = 0. Estimation: Odds-ratio for treatment effect and associated 95% CI.	ITT	P
		PP	S
	Logistic regression with treatment group and ARDS severity (severe vs mild/moderate) as fixed covariates, and investigational site as a random effect, treating participants who are lost to follow-up before Day 29 as deaths.	ITT	S
	Logistic regression with treatment group and ARDS severity (severe vs mild/moderate) as fixed covariates, and investigational site as a random effect using actual PaO_2/FiO_2 at Day 1 (as entered into eCRF).	ITT	S
	Logistic regression with treatment group, ARDS severity (severe vs mild/moderate), and gender (male or female) as fixed covariates, and investigational site as a random effect. Wald test for treatment group regression coefficient = 0.	ITT	S

The 28-day mortality rates and participants lost to follow-up in the placebo group will be compared by the placebo image assigned (5 mg BID and 15 mg BID) to determine if the rate is consistent across the 2 treatment images. Statistical comparisons between the 2 images will be performed using a Fisher's exact test in all placebo-assigned participants and within ARDS

severity subgroup. Additional endpoints may also be compared within placebo-assigned participants by treatment image.

7.4. Analysis of the Secondary Efficacy Parameters

7.4.1. In-Hospital Outcomes

The number of ventilator-free, ICU-free, oxygen-free, vasopressor-free, and hospital-free calendar days between Study Days 2 and 29 (inclusive) will be derived using the rules as outlined in Table 4. Note that qualifying vasopressor medications (ie, ATC Class C01CA - adrenergic and dopaminergic agents) will be determined through blinded review of concomitant medications, with the list of medications qualifying as vasopressors included as a note to file prior to final database lock.

Outcome-free days between Days 2 and 29 will be summarized descriptively (mean, mean, standard deviation) with treatment-group statistical comparisons performed using a Kruskal-Wallis test. Participants with missing data that do not permit evaluation for all 28 days, but are known to be alive at Day 29, will have each such day treated as not event-free. Participants who die prior to or on Day 29 will have 0 assigned days for each event-free endpoint. Participants for whom Day 29 survival status cannot be ascertained will be missing from the analysis.

Table 4: In-Hospital Outcome Event-Free Days Variable Definition

Endpoint	The Number of Study Days That a Participant Was Alive and…
Ventilator-free days	did not require mechanical ventilation.[a]
ICU-free days	is out of the ICU.[a]
Vasopressor-free days	without use of vasopressor therapy.[a]
Hospital-free days	is out of the hospital.[a]
Oxygen-free days	did not receive supplemental oxygen.[a]

[a] Participants who do not survive through Day 29 will have 0 assigned days.

7.4.2. COVID-19 9-Point Ordinal Scale

The COVID-19 9-Point Ordinal Scale (see Appendix B) will be tabulated by study day, ordinal category, and treatment group. Participants with missing values due to death will be imputed to have Category 8 (death). Participants released from the hospital that are lost to follow-up will be considered missing with regards to this endpoint. Missing intermediate days between observations will be imputed to the worst (highest) category from the 2 nonmissing observations.

7.4.2.1. Proportional Odds Model for COVID-19 9-Point Ordinal Scale

The odds of observing an improvement in the 9-point ordinal scale on Days 15 and 29 will be analyzed using a proportional odds model for the ITT and PP populations. The odds of observing a better category on the ordinal scale (lower number) will be analyzed separately at Days 15 and 29. The model will include treatment group and ARDS severity (severe vs mild/moderate) as fixed covariates and investigational site as a random effect.

As a sensitivity analysis to the proportional odds model, a Cochran-Mantel-Haenszel test for differences in treatment groups for the 9-point ordinal scale at Days 15 and 29, stratifying for ARDS severity, will also be provided.

7.4.2.2. Improvements and Changes in COVID-19 9-Point Ordinal Scale

The number and percentage of participants with a 1-point improvement at Days 15 and 29 will be tabulated and summarized by treatment group. A logistic regression model will be used to model the probability of such an improvement that includes treatment group and ARDS severity (severe vs mild/moderate) as fixed covariates and investigational site as a random effect. A similar summary will be used for 2-point improvements at Days 15 and 29.

The time from randomization to first 1- (or more) point improvement in the COVID-19 9-point ordinal scale will be derived. Participants who never improve on-study will be censored at their last assessment or Study Day 29, whichever is earlier. A Kaplan-Meier analysis will be used to describe the time to first improvement.

The mean change in the 9-point scale at Days 15 and 29 will be summarized descriptively by treatment group.

7.4.3. Sequential Organ Failure Assessment Score

7.4.3.1. Summary and Analysis of SOFA Score

The change from baseline to Day 3, 5, 8, 11, 15, and 29 in SOFA score (see Appendix C) will be summarized descriptively by treatment group.

Participants released from the hospital who are lost to follow-up will be considered missing with regards to numeric summaries of this endpoint. Missing intermediate days between observations will be imputed to the worst (highest) score from the 2 nonmissing observations.

7.4.3.2. Summary and Analysis of SOFA Score Components

Individual scoring of organ systems (respiratory, nervous system, cardiovascular, liver, coagulation, and kidneys) will be tabulated. Individual organ system scores by scheduled Study Days 3, 5, 8, 11, 15, and 29 will be compared using the Cochran-Mantel-Haenszel test stratified by baseline organ system score (0, 1, 2, 3, or 4). Participants who have died will be imputed to the worst value (4) of the component. Such comparisons will be for exploratory purposes to determine if the response is being driven by a subset of the components of the SOFA score.

7.5. Analysis of Exploratory Efficacy Variables

7.5.1. Participants Receiving IL-6, IL-6R, IL-1RA, IL-1β, GM-CSF, BTK, or JAK-Directed Therapies

The number of participants receiving IL-6, IL-6R, IL-1RA, IL-1β, GM-CSF, BTK, or JAK-directed therapies by Day 15 and by Day 29 will be summarized for the safety population. The list of such therapies will include, but is not limited to, the following: tocilizumab, canakinumab, anakinra, or sarilumab. A blinded review of concomitant medications will be performed to identify any agents meeting this criteria prior to unblinding of the clinical team.

7.5.2. Inflammatory Markers

Baseline and postbaseline levels of inflammatory markers will be summarized for the safety population. These markers include serum ferritin, CRP, D-dimer, procalcitonin, and IL-6. Changes and percentage changes in each marker will be summarized at Day 15 and Day 29.

7.5.3. SARS-CoV2 Virologic and Serologic Parameters

Baseline and postbaseline levels of viral load and anti-SARS-CoV-2 antibodies will be summarized for the safety population at Study Days 1, 15, and 29. The proportion of participants negative for SARS-CoV2 will be summarized at Day 15 and Day 29.

8. PHARMACODYNAMICS

See Section 7.5 for details regarding pharmacodynamic endpoints.

9. SAFETY AND TOLERABILITY

Appendix A provides a list of data displays. Sample data displays are included in a separate document.

9.1. General Considerations

Summary tables may be replaced with listings when appropriate. For instance, an AE frequency table may be replaced with a listing if it only contains a few unique PTs reported on relatively few participants.

9.2. Adverse Events

9.2.1. Adverse Event Definitions

A TEAE is any AE either reported for the first time or worsening of a pre-existing event after the first dose of study drug until 31 days after the last dose of study drug. Analysis of AEs (as discussed below) will be limited to TEAEs, but data listings will include all AEs regardless of their timing in relation to study drug administration. For purposes of analysis, all AEs will be considered TEAEs unless the AE can unequivocally be defined as not treatment-emergent.

Adverse events will be tabulated by MedDRA PT and SOC. Severity of AEs will be graded using the NCI CTCAE v5. The CTCAE reporting guidelines and grading details are available on the Cancer Therapy Evaluation Program website.

The subset of AEs considered by the investigator to be related to study drug will be considered to be treatment-related AEs. If the investigator does not specify the relationship of the AE to study drug, the AE will be considered to be treatment-related. The incidence of AEs and treatment-related AEs will be tabulated. In addition, serious TEAEs will also be tabulated.

9.2.2. Adverse Event Summaries

An overall summary of AEs by treatment group will include the following:

- Number (%) of participants who had any TEAEs
- Number (%) of participants who had any serious TEAEs
- Number (%) of participants who had any Grade 3 or higher TEAEs
- Number (%) of participants who had any TEAEs related to ruxolitinib/placebo
- Number (%) of participants who temporarily interrupted ruxolitinib/placebo because of TEAEs
- Number (%) of participants who permanently discontinued ruxolitinib/placebo because of TEAEs
- Number (%) of participants who had ruxolitinib/placebo dose reductions because of TEAEs
- Number (%) of participants who had any fatal TEAEs

The following summaries will be produced:

- Summary of TEAEs by MedDRA SOC and PT
- Summary of TEAEs by MedDRA PT in decreasing order of frequency
- Summary of TEAEs by MedDRA SOC, PT, and maximum severity category
- Summary of Grade 3 or higher TEAEs by MedDRA SOC and PT
- Summary of Grade 3 or higher TEAEs by MedDRA PT in decreasing order of frequency
- Summary of serious TEAEs by MedDRA SOC and PT
- Summary of serious TEAEs by MedDRA PT in decreasing order of frequency
- Summary of ruxolitinib/placebo treatment-related TEAEs by MedDRA SOC and PT
- Summary of ruxolitinib/placebo treatment-related TEAEs by MedDRA PT in decreasing order of frequency
- Summary of ruxolitinib/placebo treatment-related TEAEs by MedDRA SOC, PT, and CTCAE grade category
- Summary of Grade 3 or higher ruxolitinib/placebo treatment-related TEAEs by MedDRA SOC and PT
- Summary of ruxolitinib/placebo treatment-related serious TEAEs by MedDRA SOC and PT
- Summary of TEAEs with a fatal outcome by MedDRA SOC and PT
- Summary of TEAEs leading to ruxolitinib/placebo dose reduction by MedDRA SOC and PT

- Summary of TEAEs leading to ruxolitinib/placebo dose interruption by MedDRA SOC and PT
- Summary of TEAEs leading to discontinuation of ruxolitinib/placebo by MedDRA SOC and PT
- Summary of TEAEs requiring concomitant medications by MedDRA SOC and PT

9.3. Clinical Laboratory Tests

9.3.1. Laboratory Value Definitions

Laboratory values and change from baseline values will be summarized descriptively by visit. Baseline will be determined according to Section 4.1.3. If there are multiple values that meet the criteria for baseline, additional rules may be provided after consultation with the medical monitor to delineate which value will be defined as baseline.

Laboratory values and change from baseline values will be summarized descriptively by visit. The baseline value will be determined using the nonmissing values collected before the first dose, prioritizing scheduled assessments for baseline identification over unscheduled visits. The last record before administration in the highest priority will be considered the baseline record. For baseline laboratory candidates with the same date and time in the same priority category, additional rules may be provided after consultation with the medical monitor to delineate which value will be defined as baseline.

Laboratory test values will be assessed for severity based on the numerical component of CTCAE v5.

9.3.2. Laboratory Value Summaries

Any laboratory test results and associated normal ranges from local laboratories will be converted to SI units. Scheduled laboratory tests for hematology and clinical chemistry will be summarized by scheduled study days (eg, Days 3, 5, 7) through Day 29. Windows for postbaseline by-visit summaries will be ± 1 study day (eg, Day 3 will be Days 2-4).

When there are multiple nonmissing laboratory values for a participant's particular test within a visit window, the convention described in Table 5 will be used to determine the record used for by-visit tabulations and summaries. A record may only be used for 1 scheduled visit. If a record qualifies for more than 1 by-visit summary, then the earlier study day will take precedence.

Table 5: Identification of Records for Postbaseline By-Visit Summaries

Priority	Laboratory Visit	Proximity to Visit Window	Tiebreaker
1	Scheduled	In-window	Use smallest laboratory sequence number
2	Unscheduled	In-window	
3	Scheduled	Out-of-window	

Numeric laboratory values will be summarized descriptively in SI units, and non-numeric test values will be tabulated when necessary. In addition, line graphs and box-and-whisker plots will

be provided for hemoglobin, platelet counts, leukocytes, lymphocytes, neutrophils, creatinine, ALT, and AST.

Severity grades will be assigned to laboratory test values based on the numerical component of CTCAE v5. Shift tables will be presented showing change in CTCAE grade from baseline to worst grade postbaseline. Separate summaries for abnormally high and abnormally low laboratory values will be provided when the laboratory parameter has both high and low grading criteria. The denominator for the percentage calculation will be the number of participants in the baseline category. Shift tables will also be presented showing change in CTCAE grade from baseline to the last value using the same methods explained above. The number of participants who experienced worsening of laboratory abnormalities will be summarized by maximum severity.

9.3.3. Potential Drug-Induced Liver Injury Events

Participants with potential drug-induced liver injuries will be listed. Potential drug-induced liver injuries will be defined in participants with normal ALT and AST and total bilirubin at baseline as having elevated ALT or AST > 3 × ULN accompanied by total bilirubin > 2 × ULN at the same visit. Potential drug-induced liver injuries will be defined in participants with elevated ALT or AST or total bilirubin at baseline as having elevated ALT or AST > 2 × baseline or ALT or AST > 300 U/L, whichever occurs first, accompanied by total bilirubin > 2 × baseline and total bilirubin > 2 × ULN at the same visit.

9.4. Vital Signs

Values at each scheduled visit, change, and percentage change from baseline for vital signs, including systolic blood pressure, diastolic blood pressure, pulse, temperature, respiratory rate, pulse oximetry, and body weight will be summarized descriptively. Normal ranges for vital signs are provided in Table 6.

Table 6: Normal Ranges for Vital Sign Values

	12 – < 16 years[a]		≥ 16 years	
Parameter	**High**	**Low**	**High**	**Low**
Systolic blood pressure (mmHg)	≤ 131.0	≥ 90.0	≤ 155.0	≥ 85.0
Diastolic blood pressure (mmHg)	≤ 83.0	≥ 64.0	≤ 100.0	≥ 40.0
Pulse (beats per minute)	≤ 100.0	≥ 60.0	≤ 100.0	≥ 45.0
Body temperature (°C)	≤ 38.0	≥ 35.5	≤ 38.0	≥ 35.5
Respiratory rate (breaths per minute)	≤ 20.0	≥ 12.0	≤ 24.0	≥ 8.0

[a] Novak and Gill 2020.

10. INTERIM ANALYSES

Per SAP Amendment 3, no interim analysis will be performed.

11. CHANGES AND MODIFICATIONS TO THE ANALYSIS PLAN

All versions of the SAP are listed in Table 7.

Table 7: Statistical Analysis Plan Versions

SAP Version	Date
Original	21 APR 2020
Amendment 1	11 AUG 2020
Amendment 2	22 SEP 2020
Amendment 3	22 FEB 2020

11.1. Changes to Protocol-Defined Analyses

11.1.1. Original to Amendment 1

- The analysis for time to improvement from baseline category to earliest 1-point improvement in the ordinal scale was restated as time to improvement from baseline category to earliest improvement of any kind.
- After FDA feedback, the initial plan for treatment groups continued beyond rejection of their specific null hypothesis at the efficacy interim analysis where final reported p-values would be based on the highest level of significance achieved, either at the interim or final analysis, was modified to report p-values from the final analysis at final reporting.
- Gender was removed from analyses of the primary and secondary endpoints and added as a sensitivity analysis for the primary endpoint.

11.1.2. Amendment 1 to Amendment 2

- The weekly overall survival futility analysis was replaced by a conditional power analysis to occur after approximately every 50 participants enrolled have at least 2 weeks of follow-up.
- The reference level of ARDS severity was changed to be mild/moderate instead of severe.

11.1.3. Amendment 2 to Amendment 3

- Section 3.2, Section 3.3, Section 5.1, Section 7.3.1, Section 7.3.3, Section 10, Appendix A – The interim analysis and conditional power analysis were removed. The only remaining analysis will be the final analysis.
- Section 3.2, Section 5.3.3, Section 5.3.5, Section 7.2, Section 7.5.1, Section 7.5.2, Section 7.5.3, Appendix A – The alpha control strategy for the primary efficacy analysis was changed. The definition of the PP population was modified, and the PD-evaluable population was removed. Any planned analyses that previously used the PD-evaluable population will use the safety population instead.

11.2. Changes to the Statistical Analysis Plan

11.2.1. Original to Amendment 1

The following changes to the SAP were made based upon FDA comments on the original Protocol and SAP:

- Figure 1 – Updated to account for the changes in inclusion criteria.
- Section 7.3.1 – An estimator for the treatment group proportions and their difference was added.
- Section 7.3.2 – The age subgroup analyses were modified to < 18 years, 18-64 years, or ≥ 65 years.
- Section 7.3.3 – A tipping point analysis was added for the primary endpoint, and a comparison of the 2 placebo images (5 mg BID and 15 mg BID) was added.
- Section 7.4.2.1 – The proposed test for the proportional odds assumption was removed, and a sensitivity analysis was added to the comparison of the COVID-19 9-point ordinal scale.
- Section 7.4.3 – Additional details were added regarding the summary of SOFA components.
- Section 9.4 – Table 7 was added to define normal ranges for vital signs based on age categories.

The following editorial changes were also made based upon findings and additional feedback:

- Section 7.3.2 – Additional subgroup analyses were added for prior or concomitant remdesivir, prior or concomitant convalescent plasma, concomitant anticoagulant use, concomitant corticosteroid use, and prior biologic use.
- Section 7.4.3.1 – Editorial changes were made to the LDA model description, with minor modifications to the SAS code provided.
- Section 9.2.2 – AE summaries inconsistent with the planned tables and summaries were removed.
- Section 9.3.3 – Updated the definition of potential drug-induced liver injuries to be consistent with the Protocol.
- Section 10.3.1 – The formula for Dunnett's procedure was corrected to reflect that the joint probability should be subtracted from each marginal probability.

Other minor administrative changes have been incorporated throughout and are noted in the redline version of the document.

11.2.2. Amendment 1 to Amendment 2

The following changes to the SAP were made based upon findings and additional feedback:

- Section 4.1.4 and Section 9.2.1 – A typo was corrected for the AE follow-up period to align with the Protocol.
- Section 6.1.2 – Variables were modified and added to what will be summarized for baseline disease characteristics.
- Section 6.2 – The number and percentage of randomized participants were removed from the disposition section since this is not part of the standard disposition table, the count can be obtained from the analysis populations table for the ITT population.
- Section 7.3.3, Section 10.1, Section 10.2, and Section 10.4.2 – A Cox regression model was replaced with conditional power analysis.
- Section 7.4.3.1 – Minor modifications were made to the SAS code.

Other minor administrative changes have been incorporated throughout and are noted in the redline version of the document.

11.2.3. Amendment 2 to Amendment 3

The following changes to the SAP were made based upon findings and additional feedback:

- Section 7.4.3.1, Appendix A – The analysis of the endpoint for Change from baseline to Day 3, 5, 8, 11, 15, and 29 in SOFA score was changed from an LDA model to a simple descriptive summary.
- Section 9.5 – The electrocardiogram section was removed since that data was not collected.
- Section 6.1.3 – Summary of Prior Medications for COVID-19 was removed.

12. REFERENCES

Dunnett CW. New tables for multiple comparisons with a control. Biometrics 1964;20:482-491.

Ge M, Durham KL, Meyer DR, Xie W, Neal T. Covariate-adjusted difference in proportions from clinical trials using logistic regression and weighted risk differences. Drug Inf J 2011;45:481-493.

Novak C, Gill P. Pediatric Vital Signs Reference Chart. Peds Cases. 2018.

Pandharipande PP, Shintani AK, Hagerman HE, et al. Derivation and validation of SpO_2/FiO_2 ratio to impute for PaO_2/FiO_2 ratio in the respiratory component of the Sequential Organ Failure Assessment score. Crit Care Med 2009;37:1317-1321.

SAS Institute Inc. SAS/STAT® 13.1 User's Guide High-Performance Procedures. Cary, NC: SAS Institute Inc. 2013.

Vincent JL, de Mendonça A, Cantraine F, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. Crit Care Med 1998;26:1793-1800.

Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med 1996;22:707-710.

World Health Organization. Clinical management of severe acute respiratory infection when novel coronavirus (2019-nCoV) infection is suspected: interim guidance v. 1.2, 13 March. Geneva, Switzerland: World Health Organization; 2020.

APPENDIX A. PLANNED TABLES, FIGURES, AND LISTINGS

This appendix provides a list of the planned tables, figures, and listings for the Clinical Study Report. Shells are provided in a separate document for tables that are not in the Standard Safety Tables v1.7.

The lists of tables, figures, and listings are to be used as guidelines. Modifications of the lists that do not otherwise affect the nature of the analysis will not warrant an amendment to the SAP.

Tables

Table No.	Title	Population	Standard
Baseline and Demographic Characteristics			
1.1 Disposition			
1.1.1	Analysis Populations	ITT	X
1.1.2	Summary of Participant Disposition	ITT	X
1.1.3	Summary of Number of Participants Enrolled by Country and Site	ITT	X
1.1.4	Summary of Protocol Deviations	ITT	X
1.2 Demography and Baseline Characteristics			
1.2.1	Summary of Demographics and Baseline Characteristics	ITT	X
1.3 Baseline Disease Characteristics			
1.3.1	Summary of Baseline Disease Characteristics and History	ITT	
1.4 Prior Medication and Concomitant Medication			
1.4.1	Summary of Prior Medications	ITT	X
1.4.2	Summary of Concomitant Medications	ITT	X
1.5+ Others			
1.5.1	Summary of General Medical History	ITT	X
Efficacy			
2.1 Primary Efficacy			
2.1.1	Summary of 28-Day Mortality Analyzed by Logistic Regression Model	ITT	
2.1.2	Summary of 28-Day Mortality Analyzed by Logistic Regression Model	PP	
2.1.3	Summary of 28-Day Mortality Analyzed by Logistic Regression Model Using Actual PaO_2/FiO_2 at Randomization	ITT	
2.1.4	Summary of 28-Day Mortality Analyzed by Logistic Regression Model Treating Participants who are Lost to Follow-Up on or Before Day 29 as Deaths	ITT	
2.1.6	Summary of Tipping Point Analysis for 28-Day Mortality	ITT	
2.1.7	Summary of 28-Day Mortality for Participants Assigned Placebo by Treatment Image	ITT	
2.2 Secondary Efficacy			
2.2.1	Summary of Day 29 In-Hospital Outcomes by Treatment Group	ITT	
2.2.2.1	Summary of COVID-19 Ordinal Scale at Day 1, Day 15, and Day 29	ITT	
2.2.2.2.1	Summary of COVID-19 Proportional Odds Model at Day 15 and Day 29	ITT	

Table No.	Title	Population	Standard
2.2.2.2.2	Summary of COVID-19 Proportional Odds Model at Day 15 and Day 29	PP	
2.2.2.3.1	Summary of Participants With a 2-Point Improvement in COVID-19 Ordinal Scale on Day 15 or Day 29	ITT	
2.2.2.3.2	Summary of Participants With a 1-Point Improvement in COVID-19 Ordinal Scale on Day 15 or Day 29	ITT	
2.2.3.4	Kaplan-Meier Analysis of Time From Randomization to First Improvement in COVID-19 Ordinal Scale	ITT	
2.2.4	Summary of Change in COVID-19 Ordinal Scale at Study Day 15 and 29	ITT	
2.2.5	Summary of COVID-19 Ordinal Scale by Study Day	ITT	
2.2.6.1	Summary of Change in SOFA Score at Days 3, 5, 8, 11, 15, 29	ITT	
2.2.6.2	Summary of SOFA Score at Study Days 1, 3, 5, 8, 11, 15, and 29	ITT	
2.2.6.3	Summary of SOFA Score Components at Study Days 1, 3, 5, 8, 11, 15, and 29	ITT	
2.3 Exploratory Endpoint			
2.3.1	Summary of Participants Receiving IL-6, IL-6R, IL-1RA, IL-1 beta, GM-CSF, BTK, or JAK Directed Therapies by Day 15 and by Day 29	Safety	
2.3.2	Summary of Serum Markers of Inflammation at Baseline, Day 15, and Day 29 by Treatment Group	Safety	
2.3.3.1	Summary of SARS-CoV2 Virology at Baseline, Day 15, and Day 29	Safety	
2.3.3.2	Summary of SARS-CoV2 Antibody Testing at Baseline, Day 15, and Day 29	Safety	
Safety			
3.1 Dose Exposure			
3.1.1	Summary of Exposure and Duration of Exposure to Ruxolitinib/Placebo	Safety	X
3.1.2	Summary of Average Daily Dose by Study Day	Safety	X
3.1.3	Summary of Study Medication Compliance	Safety	X
3.2 Adverse Events			
3.2.1	Overall Summary of Treatment-Emergent Adverse Events	Safety	X
3.2.2	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.3	Summary of Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.4	Summary of Treatment-Emergent Adverse Events by MedDRA System Organ Class, Preferred Term, and Maximum Severity Category	Safety	X
3.2.6	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.7	Summary of Grade 3 or Higher Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.8	Summary of Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X

Table No.	Title	Population	Standard
3.2.9	Summary of Serious Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.10	Summary of Ruxolitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.11	Summary of Ruxolitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events by MedDRA Preferred Term in Decreasing Order of Frequency	Safety	X
3.2.13	Summary of Ruxolitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class, Preferred Term, and Maximum Severity Category	Safety	X
3.2.14	Summary of Grade 3 or Higher Ruxolitinib/Placebo Treatment-Related Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.15	Summary of Ruxolitinib/Placebo Treatment-Related Serious Treatment-Emergent Adverse Events by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.16	Summary of Treatment-Emergent Adverse Events With a Fatal Outcome by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.18	Summary of Treatment-Emergent Adverse Events Leading to Ruxolitinib/Placebo Dose Reduction by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.19	Summary of Treatment-Emergent Adverse Events Leading to Ruxolitinib/Placebo Dose Interruption by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.20	Summary of Treatment-Emergent Adverse Events Leading to Discontinuation of Ruxolitinib/Placebo by MedDRA System Organ Class and Preferred Term	Safety	X
3.2.21	Summary of Treatment-Emergent Adverse Events Requiring Concomitant Medications by MedDRA System Organ Class and Preferred Term	Safety	X
3.3 Laboratory			
3.3.1.1	Summary of Laboratory Values - Hematology	Safety	X
3.3.1.2	Summary of Laboratory Values - Chemistry	Safety	X
3.3.3.1	Shift Summary of Hematology Laboratory Values in CTC Grade - to the Worst Abnormal Value	Safety	X
3.3.3.2	Shift Summary of Chemistry Laboratory Values in CTC Grade - to the Worst Abnormal Value	Safety	X
3.3.3.3	Treatment-Emergent Worsening of Laboratory Abnormalities - Hematology	Safety	X
3.3.3.4	Treatment-Emergent Worsening of Laboratory Abnormalities - Chemistry	Safety	X
3.3.5.1	Shift Summary of Hematology Laboratory Values in CTC Grade - to the Last Value	Safety	X
3.3.5.2	Shift Summary of Chemistry Laboratory Values in CTC Grade - to the Last Value	Safety	X
3.4 Vital Signs			
3.4.1	Summary of Systolic Blood Pressure	Safety	X
3.4.2	Summary of Diastolic Blood Pressure	Safety	X

Table No.	Title	Population	Standard
3.4.3	Summary of Pulse	Safety	X
3.4.4	Summary of Respiration Rate	Safety	X
3.4.6	Summary of Weight	Safety	X
3.4.7	Summary of Pulse Oximetry	Safety	
3.4.8	Summary of Mean Arterial Pressure	Safety	

Figures

Figure No.	Title
4.1 Primary Efficacy	
4.1.2	Forest Plot of Odds Ratio and 95% CIs for 28-Day Mortality Rate
4.2 Secondary Efficacy	
4.2.1	Kaplan-Meier Estimates of Time From Randomization to First Improvement in COVID-19 Ordinal Scale
4.2.2	Stacked Bar Chart of COVID-19 Ordinal Scale by Treatment Group
4.3 Exploratory Efficacy	
4.3.1	Box-and-Whisker Plot of Ferritin by Study Visit
4.3.2	Box-and-Whisker Plot of CRP by Study Visit
4.3.3	Box-and-Whisker Plot of D-Dimer by Study Visit
4.3.4	Box-and-Whisker Plot of Procalcitonin by Study Visit
4.3.5	Box-and-Whisker Plot of IL-6 by Study Visit
4.6 Laboratory Data	
4.6.1	Box-and-Whisker Plot of Selected Laboratory Values by Study Visit

Listings

Listing No.	Title
2.1.1	Participant Disposition Status
2.1.2	Participant Enrollment and Site Information
2.2.1	Protocol Deviations
2.2.2	Randomized Participants With Inclusion/Exclusion Criteria Violations
2.3	Analysis Population
2.4.1	Demographic and Baseline Characteristics
2.4.2	Medical History
2.4.3	COVID-19 Baseline Disease Characteristics
2.4.4	Prior Medications for COVID-19
2.4.5	Prior and Concomitant Medications
2.4.6	PBRC/Platelet Transfusions
2.4.7	Procedures and Non-Drug Therapies
2.5.1	Ruxolitinib/Placebo Exposure and Compliance
2.6.1	Deaths and 28-Day Mortality
2.6.2	COVID-19 Ordinal Scale by Study Day
2.6.4	COVID-19 Improvement Events and Assessment
2.6.5	Hospitalization, Ventilation, and Intubation Status
2.6.6	SOFA Scores and Components by Study Day

Listing No.	Title
2.6.7	Arterial Blood Gas Analyses
2.6.8	Deaths
2.6.9	Chest Imaging
2.6.10	Concomitant IL-6, IL-6R, IL-1RA, IL-1 beta, GM-CSF, BTK, or JAK Directed Therapies
2.6.11	Serum Markers of Inflammation
2.6.12	SARS-CoV2 Virology
2.6.13	SARS-CoV2 Serology
2.7.1	Adverse Events
2.7.2	Serious Adverse Events
2.7.3	Grade 3 and Higher Adverse Events
2.7.4	Fatal Adverse Events
2.7.5	Ruxolitinib/Placebo Treatment-Related Adverse Events
2.7.6	Serious Ruxolitinib/Placebo Treatment-Related Adverse Events
2.7.7	Adverse Events Leading to Ruxolitinib/Placebo Dose Interruption
2.7.8	Adverse Events Leading to Ruxolitinib/Placebo Dose Reduction
2.7.9	Adverse Events Leading to Discontinuation of Ruxolitinib/Placebo
2.8.1.1	Abnormal Clinical Laboratory Values - Hematology
2.8.1.2	Clinical Laboratory Values - Hematology
2.8.1.3	Abnormal Clinical Laboratory Values - Chemistry
2.8.1.4	Clinical Laboratory Values - Chemistry
2.8.1.5	Potential Drug Induced Liver Injuries
2.8.2.1	Abnormal Vital Sign Values
2.8.2.2	Vital Sign Values

APPENDIX B. COVID-19 9-POINT ORDINAL SCALE

Participant State	Descriptor	Score
Uninfected	No clinical or virological evidence of infection	0
Ambulatory[a]	No limitation of activities	1
	Limitation of activities	2
Hospitalized Mild Disease	Hospitalized, no oxygen therapy (defined as $SpO_2 \geq 94\%$ on room air)	3
	Oxygen by mask or nasal prongs	4
Hospitalized Severe Disease	Noninvasive ventilation or high-flow oxygen	5
	Intubation and mechanical ventilation	6
	Ventilation + additional organ support – vasopressors, RRT, ECMO	7
Dead	Death	8

[a] Defined as not in hospital or in hospital and ready for discharge.
Source: WHO 2020.

APPENDIX C. SEQUENTIAL ORGAN FAILURE ASSESSMENT

The SOFA score is a scoring system to determine the extent of a person's organ function or rate of failure (Vincent et al 1996, Vincent et al 1998). The score is based on six different scores, one each for the respiratory, cardiovascular, hepatic, coagulation, renal, and neurological systems. Baseline assessment should be based on data collected closest to (but also prior to) randomization. Each subsequent day should use the worst value for each parameter in the preceding 24-hour period.

Respiratory System

When using PaO_2:

PaO_2/FiO_2 (mmHg)
- 0 = $PaO_2/FiO_2 \geq 400$
- 1 = $PaO_2/FiO_2 < 400$
- 2 = $PaO_2/FiO_2 < 300$
- 3 = $PaO_2/FiO_2 < 200$ and mechanically ventilated
- 4 = $PaO_2/FiO_2 < 100$ and mechanically ventilated

PaO_2 should be used if available. However, if PaO_2 is not available, the following table can be used to generate the SOFA score with SpO_2 (Pandharipande et al 2009):

Use of SpO_2/FiO_2 in SOFA Score

SOFA Respiratory System Points (Using SpO_2)	SpO_2/FiO_2 Ratio		
	PEEP < 8 or Not Intubated	PEEP 8-12	PEEP > 12
0	≥ 457	≥ 515	≥ 425
1	< 457	< 515	< 425
2	< 370	< 387	< 332
3	< 240	< 259	< 234
4	< 115	< 130	< 129

Note: The original SpO_2/FiO_2 Ratio as published would require a $SpO_2 > 110\%$ to achieve a SOFA score of 0. Therefore, this score has been modified to accept a SpO_2 of 96% or greater on room air as normal (ie, an SpO_2/FiO_2 ratio of ≥ 457 would be a SOFA score = 0).

Nervous System

Glasgow coma score
- 0 = GCS score of 15
- 1 = GCS score of 13-14
- 2 = GCS score of 10-12
- 3 = GCS score of 6-9
- 4 = GCS score of < 6

Cardiovascular System

Mean arterial pressure (MAP) OR administration of vasopressors required (vasopressor drug doses are in mcg/kg/min)

0 = No hypotension
1 = MAP < 70 mmHg
2 = dopamine ≤ 5 or dobutamine (any dose)
3 = dopamine > 5 OR epinephrine ≤ 0.1 OR norepinephrine ≤ 0.1
4 = dopamine > 15 OR epinephrine > 0.1 OR norepinephrine > 0.1

Liver

Total bilirubin (mg/dL)

0 = Total Bilirubin < 1.2
1 = Total Bilirubin 1.2 - 1.9
2 = Total Bilirubin 2.0 - 5.9
3 = Total Bilirubin 6.0 - 11.9
4 = Total Bilirubin ≥ 12.0

Coagulation

Platelets×103/µL

0 = Platelets ≥ 150 × 103/µL
1 = Platelets < 150 × 103/µL
2 = Platelets < 100 × 103/µL
3 = Platelets < 50 × 103/µL
4 = Platelets < 20 × 103/µL

Renal System

Creatinine (mg/dL) (or urine output)

0 = Creatinine < 1.2 mg/dL
1 = Creatinine 1.2 - 1.9 mg/dL
2 = Creatinine 2.0 - 3.4 mg/dL
3 = Creatinine 3.5 - 4.9 mg/dL, or urine output < 500 mL/d
4 = Creatinine ≥ 5.0 mg/dL, or urine output < 200 mL/d, or dialysis requirement

Source: Pandharipande et al 2009, Vincent et al 1996, Vincent et al 1998.