Efficacy and Safety of Octagam 10% Therapy in COVID-19 Patients with Severe Disease Progression

Statistical Analysis Plan

Version 3.0

Protocol Number: GAM10-10

IND Number: 022051

Protocol Version 8.0 (27 January 2021)

Sponsor: Octapharma, Inc.

Study Role	Name/Signature/Date
Octapharma [redacted]	[redacted]
CATO [redacted]	[redacted]
CATO [redacted]	[redacted]

LIST OF ABBREVIATIONS

Abbreviation	Description
AE	Adverse Event
CBC	Complete Blood Count
CMP	Comprehensive Metabolic Panel
CoV	Coronaviruses
COVID-19	Coronavirus disease-2019
CRO	Contract Research Organisation
CRP	C-reactive protein
CT	Computerized Tomography
ECMO	Extracorporeal Membrane Oxygenation
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
FAS	Full Analysis Set
FDA	Food and Drug Administration
FiO2	Fractional Inspired Oxygen
GCP	Good Clinical Practice
HFNC	High-Flow Nasal Cannula
HIT	Heparin Induced Thrombocytopenia
HIV	Human Immunodeficiency Virus
hsCRP	high-sensitivity C-reactive protein
IB	Investigators' Brochure
ICH	International Council for Harmonisation
ICU	Intensive Care Unit
IDMC	Independent Data Monitoring Committee
IEC	Independent Ethics Committee
Ig	Immunoglobin
IL	Interleukin
IMP	Investigational Medicinal Product
IMV	Invasive Mechanical Ventilation
IND	Investigational New Drug

INR	International Normalized Ratio
ITT	Intent-To-Treat
IV	Intravenous
IVIG	Intravenous Immunoglobulin
LDH	Lactic acid dehydrogenase
LOCF	Last Observation Carried Forward
MedDRA	Medical Dictionary for Regulatory Activities
MQoL-SIS	McGill Quality of Life Single-Item Scale
NCI-CTCAE	National Cancer Institute-Common Terminology Criteria for Adverse Events
NIV	Non-Invasive Mechanical Ventilation
PaO2	Arterial Oxygen Partial Pressure
PP	Per Protocol
PT	Prothrombin Time
PTT	Partial Thromboplastin Time
RNA	Ribonucleic acid
RT-PCR	Reverse Transcription Polymerase Chain Reaction
QoL	Quality of Life
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SARS	Severe Acute Respiratory Syndrome
SDV	Source Data Verification
SmPC	Summary of Product Characteristics
SOC	Standard of Care
SOP	Standard Operating Procedure

TABLE OF CONTENTS

1. Introduction

This statistical analysis plan (SAP) expands the analysis plan in *Section 9 Statistical Methods and Sample Size* of the study protocol "Efficacy and Safety of Octagam 10% Therapy in COVID-19 Patients with Severe Disease Progression" authored by the sponsor (Octapharma Inc) dated January 27, 2021. Changes to the previously approved SAP are the following:

- The interim analysis has been removed, due to an unexpected acceleration in enrollment. DMC meeting schedules have been described as per actual conduct.
- Secondary endpoints have been updated to reflect the latest protocol version.
- Missing data imputation rules have been revised to allow for the incorporation of additional subject information with respect to establishing the clinical status of a subject at each visit

The scope of analysis provided in this SAP includes: the three scheduled Independent Data Monitoring Committee (IDMC) data reviews and the final analysis after Day 33 of the last subject's enrollment.

Note that after completion of the core 33-day trial, all subjects will be followed for a one-year period by means of a registry database; subjects will be contacted every 3 months (±14 days) for their health status and any residual health effects from COVID-19 or the study treatment. Analysis of this one-year follow-up registry data is not within the scope of this SAP but will be submitted as a supplement to the core study data once all subjects have completed the year of follow-up.

2. Study Objectives and Endpoints

2.1 Primary Objective

The primary objective of this study is to determine if high-dose Octagam 10% therapy will stabilize or improve clinical status in subjects with severe COVID-19 during the first seven days following treatment commencement.

2.1.1 Primary Endpoints

- Proportion of subjects with stabilized or improved clinical status at Day 7 on at least one category on a 6-point clinical status scale.
- Change from Baseline (Day 1) at Day 7 in terms of the 6-point clinical status scale (descriptive analysis).

Clinical status categories will be defined as:

1. Hospital discharge or meet discharge criteria (discharge criteria are defined as clinical recovery, i.e., no fever, respiratory rate, oxygen saturation return to normal, and cough relief).
2. Hospitalization, not requiring supplemental oxygen.
3. Hospitalization, requiring supplemental oxygen (but not non-invasive mechanical ventilation (NIV)/high-flow nasal cannula [HFNC]) as defined by A-a Gradient < 150 mmHg.
4. Intensive Care Unit (ICU)/hospitalization, requiring NIV/HFNC therapy, as defined by A-a Gradient ≥ 150 mmHg.
5. ICU, requiring extracorporeal membrane oxygenation (ECMO) and/or IMV.
6. Death.

A subject meets the success criteria of stabilization or improvement if the clinical status category on Day 7 is not worse than the clinical status category at Baseline (Day 1 prior to IMP infusion).

2.2 Secondary Objectives

Key Secondary Efficacy Objectives

- To compare the length of hospital stay in subjects treated with *Octagam 10%* compared to those that received placebo through Day 33.
- To determine if high-dose *Octagam 10%* therapy will stabilize or improve clinical status in subjects with severe COVID-19 during the 14 days following treatment commencement.
- To compare the cumulative duration of IMV in subjects treated with *Octagam 10%* compared to those that received placebo through Day 33.
- To compare the cumulative incidence of severe progression in subjects treated with *Octagam 10%* compared to those that received placebo at Day 33.
- To compare cumulative duration of time in ICU in subjects treated with *Octagam 10%* compared to those that received placebo at Day 33.
- To compare cumulative mortality in subjects treated with *Octagam 10%* compared to those that received placebo at Day 33.

Other Secondary Efficacy Objectives

- To determine if high-dose *Octagam 10%* affects severe acute respiratory syndrome from Coronavirus 2 (SARS-CoV-2) infection status at Day 7.

- To determine if high-dose *Octagam 10%* therapy will improve clinical status in subjects with severe COVID-19 during the 14 days following treatment commencement.
- To compare the cumulative incidence of IMV in subjects treated with *Octagam 10%* compared to those that received placebo through Day 33.
- To compare the time to recovery in subjects treated with *Octagam 10%* compared to those that received placebo through Day 33.
- To compare the time to first improvement in clinical status in subjects treated with *Octagam 10%* compared to those that received placebo through Day 33.

Safety Objectives

- To describe the safety of Octagam 10% compared to placebo when administered to subjects with severe COVID-19.

2.2.1 Secondary Endpoints

Key Secondary Efficacy Endpoints

- Length of hospital stay (time to discharge) through Day 33
- Proportion of subjects with maintenance or improvement by at least one category on the 6-point clinical status scale on Day 14.
- Cumulative duration of IMV through Day 33
- Proportion of subjects with severe progression defined as requiring ECMO, IMV and/or have died through Day 33
- Length of time in ICU from randomization through Day 33
- Mortality rate through Day 33.

Other Secondary Efficacy Endpoints

- Results of reverse transcription polymerase chain reaction (RT-PCR) for SARS-CoV-2 from nares/throat swab and/or sputum and/or lower respiratory tract sample on Day 7.
- Proportion of subjects admitted to ICU from randomization through Day 33
- Proportion of subjects with improvement on Days 7, 14, 21, 33
- Proportion of subjects requiring IMV by Day 33.
- Time to recovery through Day 33 where recovery is defined as a clinical status of 1 or 2

- Time to first improvement in clinical status through Day 33

Safety Endpoints

- Incidence of all adverse events (AEs)
- Incidence of AEs considered related to the Investigational Medicinal Product (IMP)
- Incidence of serious adverse events (SAEs)
- Clinical laboratory parameters
- Vital sign parameters
- Radiological findings (chest computerized tomography [CT]/chest X-ray)

2.3 Exploratory Objectives

- To compare changes in SpO2 on room air from Baseline (Day 1) to all available on-study time points up to Day 33 in subjects treated with Octagam 10% compared to those that received placebo.
- To compare changes in the Modified Borg Dyspnea Scale from Baseline (Day 1) to Days 7, 14, 21, and 33 in subjects treated with Octagam 10% compared to those that received placebo.
- To compare changes in Quality of Life (McGill Quality of Life Single-Item Scale, MQoL-SIS) score from Baseline (Day 1) to Days 7, 14, 21 and 33 in subjects treated with Octagam 10% compared to those that received placebo.

2.3.1 Exploratory Endpoints

- Improvement of SpO_2 on room air from Baseline (Day 1) to Day 33 using trend from all available observed data.
- Modified Borg Dyspnea Scale score at Baseline (Day 1) and on Days 7, 14, 21, and 33.
- MQoL-SIS score at Baseline (Day 1) and on Days 7, 14, 21, and 33.

3. Study Design, Procedures, and Analysis Timing

This is a randomized, double-blind, placebo-controlled, multicenter, Phase 3 study in the US, Russia and Ukraine to evaluate whether high-dose Octagam 10% therapy can stabilize or improve clinical status in subjects with severe COVID-19 disease.

Following screening procedures, eligible subjects will be randomized in a 1:1 ratio to receive either Octagam 10% in a dose of 2 g/kg over 4 days (0.5 g/kg/day), or an equivalent volume of saline (placebo).

The infusion will be administered over a period of 2 hours per day over a total of 4 consecutive days. The first dosing will take place within 24 hours of the Screening Visit.

Approximately 208 male or female adult subjects will be recruited so that approximately 104 subjects in each treatment arm will complete the study. Each subject's participation will last a total of approximately 33 days, including 29 days for safety follow-up, followed by a one-year registry follow-up.

The final primary analysis will take place after database lock.

The primary analysis may also be conducted prior to database lock but after the last subject enrolled completes their Day 33 visit (or discontinues). If conducted, this analysis would be distributed to a limited unblinded team to evaluate the need to submit an emergency use application. Full details on this process will be provided in an unblinding plan and the clinical study report.

Upon end of study follow up, an end of study analysis may be conducted.

4. General Analytic Considerations

Unless otherwise mentioned:

- Number of subjects in the defined population and number of subjects with available data will be provided.
- Continuous variables will be summarized using descriptive statistics such as mean, standard deviation, median, quartiles, minimum, and maximum, and 95% confidence interval.
- Categorical variables (e.g., gender, race, etc.) will be provided using frequency tabulations.
- Individual listings will be provided to support the descriptive analyses.
- Baseline is defined as the last observation prior to randomization for the analysis of efficacy endpoints. Baseline is defined as the last observation prior to first administration of study drug for the analysis of safety endpoints.

4.1 Missing Data

The latest patient status available prior to Day 7 will be used to determine clinical status in the absence of a clinical status assessment. If a death date is prior to Day 7, clinical status at Day 7 will be identified as a 6. Otherwise, if a subject's last status reflects that they were in the ICU and receiving ECMO and/or IMV, clinical status will be imputed as a 5. Otherwise, if a subject's last status reflects that they were in the ICU and not receiving ECMO and/or IMV, clinical status will be imputed as a 4. If the LOCF clinical status value for a patient is higher than these imputed values or if a subject status cannot be derived based on available information on ECMO, IMV, or ICU status, the last observation carried forward (LOCF)

will be used to impute patient values. Subjects with no post-baseline clinical status assessment, or prohibited medication use, on or prior to Day 7 will be imputed as failure.

An imputation analogous to the primary endpoint will be performed for the secondary endpoint of clinical status at Day 14.

For the secondary endpoint length of hospitalization, Subjects who are still hospitalized as of Day 33, subjects with missing end dates due to withdrawal, subjects with unusable end dates due to prohibited medication use, or those who died on/prior to Day 33 will be censored at Day 33. Discharge occurring on or after the start date of prohibited medication use will not be considered an event.

Cumulative duration in ICU will be imputed similarly, unless a patient has a useable discharge date (e.g. not occurring prior to prohibited medication use); in this case, the ICU discharge date will be set to the hospital discharge date. If a subject's last assessment identifies that IMV was used (and has no usable end date), the date on which it is no longer used will be assumed Day 33 unless the subject is discharged from the hospital prior to Day 33. If this occurs, the hospital discharge date will be used. IMV end dates will be imputed similarly, assuming that IMV ceases at the time of ICU discharge. Hospitalization, ICU, and IMV end dates for subjects who died prior to Day 33 during these events (e.g. after ICU admission or IMV initiation) will result in assumed end dates of Day 33 for endpoint analyses, so long as prohibited medication was not administered. Any dates on or after the start date of prohibited medication use will be considered unusable for imputation purposes; ongoing hospitalization, ICU, or IMV at the time of medication use will default to extrapolation to Day 33.

For the key secondary endpoint involving severe progression (ECMO, IMV, and death), subjects who withdrew early from study prior to Day 33 without evidence of events will be treated as not having the event. Sensitivity analyses will be conducted (1) treating them as failures (having the event at their last study visit) and (2) excluding them from the analyses. An additional analysis of these endpoints will be provided, treating prohibited medication use prior to Day 33 as an event, as well as allowing for status (event or success) without regard to prohibited medication use.

The outcomes for safety assessment and for additional efficacy analysis as well as exploratory endpoints will not be imputed if data is missing.

Please also refer to the more detailed descriptions in sections 12 to 18 of this SAP for additional details on the handling of missing data for particular endpoints and analyses.

4.2 Analysis Scope of Independent Data Monitoring Committee (IDMC) Meeting and Final Analysis

The analysis plan for the three scheduled IDMC data reviews can be found in Section 8 of this SAP. The rest of the analysis plan applies to the final analysis that will occur after all subjects have reached Day 33.

4.3 Covariate Adjustment

Testing will be stratified by age category (≤ 65 versus > 65).

4.4 Test Size and Confidence Levels

Unless otherwise noted, all reported confidence intervals will be computed at the 95% coverage level, and all p-values will be assessed at the two-sided 0.05 significance level. Control for multiplicity is only planned for the analysis of secondary endpoints using a hierarchical testing approach.

4.5 Analysis of Safety Data

If not further specified, the system organ class (SOC) and preferred term for adverse events analysis will be summarized using Medical Dictionary for Regulatory Activities (MedDRA) Version 23.0 with severity grade defined by CTCAE Version 5.0.

4.6 Reporting Timeframe of Measurement for Analysis

For analyses of primary and secondary efficacy endpoints, including clinical status and RT-PCR, the latest observation within any given study day will be used. If the observations for the Day 7 visit falls outside of the 24-hour window of the study day or the Day 14 visit falls outside of the 48-hour window of the study day, the observation will be considered as missing and the imputation described in Section 4.1 will apply as appropriate. Otherwise, unless specifically stated otherwise, analyses will be conducted based on nominal visits.

For summary analyses of efficacy assessment, including oxygenation, oxygen saturation, modified Borg Dyspnea Scale and McGill Quality of Life Single-Item Scale (MQoL-SIS) and safety parameters (including vital signs, clinical laboratory evaluations and chest CT/chest X-ray findings), the latest observation within the study day during treatment period will be used. No imputation will be made for these assessments.

For safety analysis with adverse events, the observation with highest severity within the study day during treatment period will be used. No imputation will be made.

5. Study Size and Power

Assuming that the proportions of subjects who successfully achieve primary endpoint are 0.6 for placebo arm and 0.8 for IVIG arm, with additional assumptions of normality approximation of difference of two proportions, the maximum sample size needed is 208 to provide at least 80% power assuming a drop-out rate of 15%. That is, 176 subjects with complete data for primary endpoint (clinical status at baseline and Day 7) are needed to maintain at least 80% of statistical power under overall type I error of 0.05.

6. Analysis Populations

In the definitions below, an Analysis Population refers to a particular set of participants. Additional censoring rules for analyses based on these populations are described in relevant sections of the SAP. The definition applies to analyses for scheduled IDMC meetings and final analysis.

- Screened Population – The Screened Population includes all subjects who are screened.
- Enrolled Population – The Enrolled Population includes all subjects who enrolled in the study.
- Intent-to-Treat (ITT) Population – The Intent-to-Treat Population includes all subjects who are randomized.; Analysis will be conducted according to the study arm in which subjects are randomized, regardless of randomization errors or protocol violations.
- Safety (SAF) Population – The Safety Population includes all subjects who are randomized and received any amount of IVIG or placebo. The analysis will be conducted according to the treatment which subjects actually received. This analysis population is a subset of the ITT population.
- Per-Protocol (PP) Population - The Per-Protocol Population includes subjects who receive at least 75% of the planned dose (3 completed infusions of IVIG or placebo) without major protocol deviations that would potentially interfere with the efficacy assessment. This will be determined in consultation with the Octapharma Study leadership during blinded data review. The criteria for exclusion of subjects from the PP Population will be determined and documented prior to scheduled unblinding. The analysis will be conducted according to the treatment which subjects actually received. This analysis population is a subset of the ITT population.

Summary of Planned Analysis, by Population

Analysis	Screened	EP	ITT	Safety	PP
Disposition	√				
Subject Assessment		√			
Study Treatment Dosage			√	√	√
Primary Efficacy			√		√
Secondary Endpoints			√		√
Exploratory Endpoints			√		
Safety				√	
Additional Efficacy			√	√	√
Additional Safety				√	
Concomitant Medications and Procedures				√	

7. Blinded Data Review

Prior to scheduled study unblinding, the data will be reviewed in a blinded fashion for final data analysis decisions. Based on analysis needs, the lead statistician will develop specific listings for this review and will document decisions made prior to unblinding. These decisions may include (but not necessarily be restricted to) protocol violations leading to inclusion/exclusion of subjects in the different analysis populations and/or data points in specific analyses. More specifically, protocol violations will include those documented in the protocol violation CRF as well as violations that can be detected directly from the data base. Decisions on missing data, identification of prohibited medications and final imputation of endpoints will be done during this blinded review phase. Blinded data review may also include clinical evaluation of safety data for reasonableness or to issue clinical queries to the site if needed. Any decisions affecting the analysis will be described in internal documentation and the clinical study report.

8. Independent Data Monitoring Committee (IDMC) Meeting Scope and Time Point

Due to accelerated enrollment, there have been two scheduled IDMC meetings to review interim data (i.e., not including operational meetings) for safety monitoring purposes. The time points specified below were the timepoints for the scheduled IDMC meetings, and the corresponding analysis scope is specified.

1. Approximately 44 subjects (first 25% of data collected) for safety review

The analysis scope of this IDMC review focuses on safety data review. The blinded study statistician prepared overall data presentations for the study team, whereas the dedicated unblinded statistician prepared parallel data presentations by treatment assignments and provided unblinded data to the IDMC only.

Proportions of subjects screened, enrolled, randomized, and length of follow-up were provided by site for the Screened Population. Screening failure reasons were summarized. Furthermore, subject disposition based on the Full Analysis Population, including subject status. Dose administration status overall and by treatment arm was provided for the Safety Population.

Baseline information, demographics, treatment status, primary reason for early termination, and baseline laboratory values were tabulated and summarized for the Safety Population. In addition, adverse events (AEs), serious adverse events (SAEs), treatment-emergent adverse events (TEAEs), and TEAEs related to study drug by maximum severity was summarized by System Organ Class (SOC) and Preferred Term (PT) by CTCAE grades.

Demographic information, screening failure reasons, protocol deviations, subject treatment status, reasons for early termination, and adverse events were listed.

2. An additional data review at 50% of subjects enrolled, based on similar reports to those specified above, without a formal meeting.

Further details with respect to IDMC decision reviews and decision making can be found in the IDMC Charter and meeting minutes. An additional meeting is planned after 100% of data is accumulated.

9. Subject Disposition

The frequency and percentage of subjects screened, enrolled, randomized, as well as screening failure reasons will be tabulated based on the Screened Population. The number of subjects enrolled and randomized at each site will be tabulated by treatment arm using the ITT Population.

Final status, including frequency and percentage of subjects who completed the study, were lost to follow-up, or terminated early along with primary reasons for discontinuation will be tabulated by treatment arm for the ITT Population.

The frequency of protocol deviations will be summarized by treatment arm for the ITT Population. Furthermore, all protocol deviations classified as major and/or significant with respect to the evaluation of the primary endpoints will be listed separately.

10. Summary of Subject Assessment

The following assessments will be summarized:

- Demographics and Baseline Characteristics (ITT population, Safety Population, and Per Protocol Population)

  Baseline demographic and characteristics (e.g. age, gender, height, weight, race, ethnicity)
- Medical History (ITT population, Safety Population, and Per Protocol Population)

  Medical history data by Medical Dictionary for Regulatory Activities (MedDRA) System Organ Class and Preferred Terms.
- Physical Examinations, Vital Signs, Weight, Oxygen Saturation (SpO2), Laboratory Results (Safety Population)

  Physical exam, vital signs, weight, SpO2, spirometry and laboratory results including chemistry and hematology, will be summarized by visit.
- Clinical Status and Oxygen/Ventilator Status (ITT population and Safety Population)

  The clinical status and ventilator status will be tabulated according to the categories defined in the protocol by visit, regardless of prohibited medication use, based on shifts from baseline displays. Additional summaries may be provided, suppressing information following prohibited medication use, if use of prohibited medications is highly prevalent.
- rRT-PCR (ITT population)

  The result of rRT-PCR will be tabulated according to predefined categories by visit and sample type.
- Chest X-ray (Safety Population)

  The assessment of pulmonary infiltrate and pleural effusion over time will be tabulated separately for pre-specified age strata and visit.

- Borg Dyspnea Scale and McGill Quality of Life Single Item Scale (ITT Population)

  The scale measures will be summarized by descriptive statistics by visit.

11. Analysis of Study Treatment Dosage

Treatment dosage will be presented descriptively by study day and treatment, including but not limited to number of planned and actual doses and number of doses stopped or interrupted. All planned and actual doses will be listed as well.

The analysis will be presented for the ITT Population and repeated for the Safety and Per-Protocol Populations. Listings will be provided to support the descriptive analysis.

12. Definition of Study Outcome and Endpoint

This section includes definition of metrics for analysis specified under the scope of this SAP. The handling of missing data, including data collected after prohibited medication use, is detailed in the corresponding analysis section for each metric.

Stabilized or Improved Clinical Status / Maintenance or Improvement on the Clinical Status

A subject's clinical status is considered stable or improved if the clinical status category is the same or a better clinical category from baseline to the study day of interest.

Proportion of subjects with maintenance or improvement by at least one category on the 6-point clinical status scale

This is defined as the proportion of subjects who were identified as stable (same clinical status as baseline) or improved (at least 1 level decrease in clinical status score from baseline) at the visit day.

Proportion of subjects with improvement by at least one category on the 6-point clinical status scale

This is defined as the proportion of subjects who were identified as improved (at least 1 level decrease in clinical status score from baseline) at the specified visit day.

Time to recovery through Day 33 where recovery is defined as a clinical status of 1 or 2

Time to recovery is defined as (first date of clinical status 1 or 2 – randomization date + 1). Subjects which do not have this event, or who are considered failures based on clinical status imputation methods corresponding to the primary analysis, will be considered censored at Day 33.

Time to first improvement through Day 33 where recovery is defined as a clinical status of 1 or 2

Time to recovery is defined as (first date of clinical status improvement [defined as a 1-point or more decrease in scale from baseline] – randomization date + 1). Subjects which do not have this event, or who are considered failures based on clinical status imputation methods corresponding to the primary analysis, will be considered censored at Day 33.

Length of hospital stay (time to discharge) through Day 33

Length of hospitalization is defined as (date of discharge [event] – randomization date + 1). Subjects which do not have this event will be considered censored at Day 33. Only the index hospitalization will contribute to this endpoint; readmittance will not be considered for analysis, but these incidences will be provided in data listing. Additional event (hospital discharge) and censoring (no hospital discharge) details are included in a subsequent section.

Cumulative duration of IMV through Day 33

Cumulative duration on IMV through Day 33 and is defined as (IMV end date – IMV start date + 1). If a subject was repeatedly put on IMV on the same day, the study day will only be counted once. If a subject was placed on IMV more than once, the cumulative duration will reflect the total number of days on ventilation (e.g. sum of each IMV end date – IMV start date + 1).

Proportion of subjects requiring IMV through Day 33

This is defined as the proportion of subjects who had IMV in the period from date of randomization to Day 33. If a subject has two or more events during this time, (s)he will still be counted only once.

Proportion of subjects with severe progression defined as requiring ECMO, IMV and/or have died through Day 33

This is defined as the proportion of subjects who had any of these progressive events in the period from the date of randomization to Day 33. If a subject has two or more events during this time, (s)he will still be counted only once.

Length of time in ICU through Day 33

Length of time in ICU includes number of days spent in the ICU from randomization through Day 33 and is defined as [ICU discharge date – (later of ICU admittance date, randomization date) + 1]. If a subject was discharged and readmitted to the ICU on the same day after randomization, the study day will only be counted once. If a subject was admitted to the ICU more than once, the length of time in ICU will reflect the total number of days on ventilation [e.g. sum of each ICU discharge date – (later of IMV admittance date, randomization date) + 1]).

Proportion of subjects admitted to the ICU from randomization through Day 33

This is defined as the proportion of subjects who were admitted to the ICU in the period from the date of randomization to Day 33. If a subject has two or more events during this time, (s)he will still be counted only once.

Mortality rate through Day 33

This is defined as the proportion of subjects who died in the period from the date of randomization to Day 33.

Proportion of Subjects who are Tested as "Not Detected" on Day 7

This is defined as the proportion of subjects with "not detected" result for SARS-CoV-2 from an rRT-PCR test on Day 7. If the subject has multiple measurements on the same study day, the last measurement will be used.

13. Analysis of Primary Efficacy Endpoint

The primary analysis will be based on the ITT Population. The null hypothesis of no difference in proportion of subjects who reach success (clinical status not worse than the clinical status category at Baseline at Day 7) in two treatment groups will be tested versus the alternative hypothesis that there is a difference between proportions of subjects who reach success in two groups. Hypothesis testing will be performed using Cochran-Mantel-Haenszel (CMH) stratified by age category (≤65 versus >65) at a two-sided 0.05 significance level. The difference in proportions between two groups with standard deviation, and 95% Miettinen-Nurminen (1985) confidence interval (CI) will also be provided.

If the CMH fails the Mantel-Fleiss criteria (1980), a chi-squared test will be the basis of evaluation and unweighted asymptotic CIs for the difference in proportions will be used. If, when implementing a chi-squared test, an expected cell count < 5, Fisher's exact test will be used and exact (Santner-Snell, 1980) CIs will be provided.

For subjects with clinical status assessment at Baseline (Day 1), missing data at Day 7 will be imputed based on the method described in Section 4.1. Subjects with missing data at Baseline (Day 1) will be imputed as failure.

Subjects who use prohibited medication (medications being administered as part of other experimental studies or the use of experimental non-IVIG products, COVID-19 convalescent plasma, or interferons) on or before Day 7 will be counted as failure. Prohibited medications will be identified during blinded review described in Section 7.

Shifts from Baseline (Day 1) at Day 7 in terms of the 6-point clinical status scale will be reported (descriptive analysis).

The durability of the treatment effect will be reviewed based on the comparison of proportions who are stabilized or improved between treatment arms at Day 14, Day 21, and Day 33 in the final analysis. No formal hypothesis testing will be conducted for the Day 21 and Day 33 analyses and the descriptive comparison will be carried out based on the method applied for the primary analysis. Graphical representations of clinical status, and changes from baseline, over time will be presented by treatment group.

Sensitivity Analysis

The primary endpoint analysis will be repeated using the PP and Safety Populations as sensitivity analyses.

A complete case analysis, using patients with completed clinical status assessments at Day 7, will be presented.

An analysis of the primary endpoint, using all data collected regardless of prohibited medication use, will also be conducted.

A worst-case scenario will be conducted as an additional sensitivity analysis, counting all subjects with missing data at Baseline (Day 1) or Day 7 as failures.

14. Analysis of Secondary Endpoints

The analyses for secondary endpoints will be based on the ITT Population. The hypothesis testing with p-value assessments for secondary endpoints at designated timepoints will be implemented for confirmatory purpose after statistical significance is claimed from the primary endpoint. The following secondary endpoints will be analyzed by hierarchical procedure in the order specified below with overall 2-sided type I error of 0.05:

1. Length of hospital stay (time to discharge) through Day 33
2. Proportion of subjects with maintenance or improvement by at least one category on the 6-point clinical status scale on Day 14.
3. Cumulative duration of IMV through Day 33
4. Proportion of subjects with severe progression defined as requiring ECMO, IMV and/or have died through Day 33
5. Length of time in ICU from randomization through Day 33
6. Mortality rate through Day 33

The hierarchical testing procedure will stop when one of the endpoints cannot demonstrate statistical significance, i.e. as soon as the first statistical test yields a p-value ≥ 0.05. If this happens, the subsequent endpoints will not be tested.

Length of hospitalization is defined as (date of discharge – randomization date + 1). Subjects who are still hospitalized as of Day 33, subjects with missing end dates due to withdrawal, subjects with unusable end dates due to prohibited medication use, or those who died on/prior to Day 33 will be censored at Day 33. The median time to discharge and associated 95% CI for each treatment arm will be estimated using the Kaplan-Meier method. Comparisons between treatment arms will be performed using a log-rank test, stratified by age category. A rate ratio and associated 95% CI, equivalent to a hazard ratio from a Cox proportional hazard model stratified by age, will also be presented.

Proportion of subjects that stabilized or improved at Day 14 is defined as per the primary endpoint and hypothesis testing for proportions will be conducted as described for the in the primary endpoint section.

Cumulative duration on IMV through Day 33 and is defined as (IMV end date – IMV start date + 1). Subjects who do not require IMV will have a duration set to 0. Subjects who die on or prior to Day 33 or are still on IMV at Day 33, will have a duration set to their actual/projected Day 33 date for the IMV end date, so long as prohibited medication was not administered. If IMV end date is missing but subject is discharged from ICU or hospital, then IMV end date will be set to the earlier of the ICU or hospital discharge date. Otherwise, subjects without usable IMV end date information due to withdrawal or prohibited medication use will use a projected Day 33 date for the IMV end date. Comparisons between

treatment arms will be performed using a stratified Wilcoxon-Mann-Whitney (van Elteren) test with age category as a stratification factor.

Proportion of subjects with severe progression will be defined as the proportion that require ECMO, IMV or died prior to Day 33. For this analysis, subjects who use prohibited medication prior to the earliest of progression or Day 33 will not be included in analyses. Hypothesis testing for this endpoint will be conducted as described for the primary endpoint.

Length of time in ICU from randomization through Day 33 includes number of days spent in the ICU from randomization through Day 33 and is defined as [ICU discharge date – (later of ICU admittance date, randomization date) + 1]. Subjects who are not admitted to the ICU will have a duration set to 0. Subjects who die on or prior to Day 33 or are still in the ICU at Day 33, will have a duration set to their actual/projected Day 33 date for the discharge date, so long as prohibited medication was not administered. If ICU discharge date is missing but subject is discharged from hospital, then ICU discharge date will be set to hospital discharge date. Otherwise, subjects without usable ICU end date information due to withdrawal or prohibited medication use will use a projected Day 33 date for the ICU end date. Comparisons between treatment arms will be performed using a stratified Wilcoxon-Mann-Whitney (van Elteren) test with age category and baseline ICU status (In ICU at baseline versus not in ICU at baseline) as stratification factors.

Mortality rate is defined as the proportion of subjects who have died on or prior to Day 33. For this analysis, subjects who use prohibited medication prior to the earliest of death or Day 33 will not be included in analyses. Hypothesis testing will be performed as described for the primary endpoint.

Key secondary endpoints which demonstrate statistical significance will be analyzed by subgroups including sex (male/female), race, ethnicity, geographic region, and disease severity (according to the definitions in the eCRF). All key secondary endpoints will be analyzed by age using testing analogous (unstratified) testing methodology to the specified analyses. Length of time in ICU from randomization through Day 33 will also be analyzed by baseline ICU status. Sensitivity analyses, using all available data, regardless of prohibited medication usage, will be conducted where applicable. The descriptive analyses for the following, other secondary endpoints will be based on the primary efficacy analysis (e.g. CMH) method:

- Proportion of patients whose results of reverse transcription polymerase chain reaction (RT-PCR) for SARS-CoV-2 from nares/throat swab and/or sputum and/or lower respiratory tract sample are “Not Detected” on Day 7
- Proportion of subjects admitted to the ICU from randomization through Day 33
    - Only admittance (new or returning) post-randomization will qualify as an event
    - Analyses will be stratified additionally by baseline ICU status
    - If expected event counts do not satisfy CMH testing criteria, baseline ICU status will be removed from analysis as a first step in analysis update
- Proportion of subjects with improvement on Days 7, 14, 21, 33

- Proportion of subjects requiring IMV by Day 33.

The descriptive analyses for the following, other secondary endpoints will be based on the Kaplan-Meier analysis of length of hospitalization:

- Time to recovery through Day 33 where recovery is defined as a clinical status of 1 or 2
- Time to first improvement in clinical status through Day 33

These secondary endpoints will also be analyzed by age category.

15. Analysis of Exploratory Endpoints

The analysis will be based on the ITT Population corresponding to each outcome. Missing data will not be imputed and endpoints will not be tested statistically. Endpoints will be presented descriptively by treatment arm.

- **Oxygen saturation on room air**

The associated endpoint is "trend of improvement of SpO2 on room air from Baseline (Day 1) to Day 33." A side-by-side box plot of SpO2 level by treatment group will be provided.

- **Modified Borg Dyspnea Scale Score at Baseline (Day 1) and on Days 7, 14, 21, and 33**

The associated endpoints are "changes in modified Borg Dyspnea Scale from baseline (Day 1) to Days 7, 14, 21 and 33." Descriptive statistics for crude value and absolute changes from baseline (e.g. Day 7 measurement minus baseline measurement) by treatment arm and overall will be provided for the ITT Population. Moreover, descriptive statistics for difference in change from baseline between the treatment arms will be provided for each time point. Descriptive statistics, least squares means, standard errors, and 95% CIs for mean differences will be presented. Least squares estimates presented at each visit will be based on ANCOVA models with baseline value and age category as covariates.

- **MQoL-SIS score at Baseline (Day 1) and on Days 7, 14, 21 and 33**

The associated endpoints are "changes in MQoL-SIS score from baseline (Day 1) to Days 7, 14, 21 and 33." Descriptive statistics for crude value and absolute changes from baseline (e.g. Day 7 measurement minus baseline measurement) by treatment arm and overall will be provided for the ITT Population. Moreover, descriptive statistics for difference in change from baseline between the treatment arms will be provided for each time point. Descriptive statistics, least squares means, standard errors, and 95% CIs

for mean differences will be presented. Least squares estimates presented at each visit will be based on ANCOVA models with baseline value and age category as covariates.

16. Safety Analysis

This analysis will be based on the Safety Analysis Population and will be presented by treatment arms as well as overall.

All AEs will be tabulated by severity from baseline up to Day 33 after first dose.

All AEs will be coded using the MedDRA dictionary Version 23.0 or higher. The severity of AEs will be graded 1 to 5 according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) Version 5.0 . AEs with missing severity will be presented in the summary table as CTCAE Grade 3.

A treatment-emergent adverse event (TEAE) is defined as an AE which started on or after the first dosing date of the study drug or an existing condition becoming worse on or after the first dosing date of the study drug. Adverse events with unknown onset date and with end date after first dosing date will be counted as a TEAE, unless it can be definitely shown that the condition was already present before the first dosing.

A treatment-related AE is defined as an AE which has a suspected relationship to the study drug, i.e. which has been rated as possibly or probably related by the investigator. Adverse events with missing relationship to study drug will be assessed as treatment-related.

The following summaries will be provided:

- A table for overall summary and a listing of all AEs that occurred during the screening period by SOC and PT will be provided.
- The following summary tables will be generated for AEs occurring from randomization date to Day 33:
- Overall summary
- All TEAEs
- TEAEs leading to discontinuation of treatment

- TEAEs with toxicity grade >=3
- Drug-related TEAEs
- Serious TEAEs
- TEAEs that led to death

Unless otherwise specified, AEs will be summarized by treatment group, SOC and PT. All TEAEs will be summarized by maximum toxicity grade, SOC, and PT.

If a subject experiences multiple AEs in the same SOC, that subject will be counted only once for that SOC. If a subject experiences multiple AEs under the same PT, then the subject will be counted only once for that PT. If a subject experiences the same AE more than once with different toxicity grades, the event with the highest grade will be tabulated in maximum toxicity grade tables.

Absolute change in vital signs and clinical laboratory parameters from baseline to each study visit up to Day 33 will be summarized descriptively with mean, standard deviation, median, range and 95% confidence intervals. The clinically significant status of vital signs will also be tabulated.

No missing data will be imputed, and no formal hypothesis testing will be performed for safety endpoints.

Subject level listings of AEs, vital signs, clinical laboratory parameters and inflammation assessment will be provided.

17. Additional Efficacy Analysis

The proportions of subjects stabilized or improved will be provided at Day 21 and Day 33 for both treatment arms. No hypothesis testing will be made for this analysis, but the difference of proportions will be presented, together with its standard error and the associated 95% confidence interval.

The proportions and frequencies of subjects' clinical status will be summarized over time. Frequencies and proportions of oxygen ventilator status will be also summarized over time. No hypothesis testing will be made for this analysis, but the difference of proportions will be presented, together with its standard error and associated 95% exact confidence interval.

18. Additional Safety Analysis

Additional to the analysis for adverse events, further safety data will be analyzed as described in this section. This analysis will be based on the Safety Population and will be presented by treatment arms as well as overall. No missing data will be imputed.

18.1 Analysis of Prior and Concomitant Medications

Prior medications are defined as medications taken within 1 week of the Screening Visit until the start of IMP infusion. Concomitant medications are defined as medications with a start date and time after the start of the Day 1 infusion.

Frequency summaries of prior and concomitant medications coded with WHO drug dictionary (version Mar2020) will be provided by treatment group. Medications may also be further categorized by grouped coded terms for summary.

At each level of subject summarization, a subject is counted once if the subject reported one or more medications. A listing, including reason for use, duration, frequency and dosage of concomitant and prior medication will be provided.

18.2 Analysis of Prior and Concomitant Procedures

Prior procedures are defined as procedures with a start date and time before the start of the Day 1 infusion. Concomitant procedures are defined as procedures with a start date and time after the start of Day 1 infusion.

Frequency summaries of prior and concomitant procedures and surgeries coded using MedDRA (version 23.0) will be provided by treatment group. At each level of subject summarization, a subject is counted once if the subject reported one or more procedures. A listing, including necessity of surgery and duration of concomitant and prior procedures will be provided.

18.3 Analysis of Additional Safety Data

Descriptive statistics of changes from baseline to each scheduled post-baseline assessment will be provided for vital signs, laboratory test results, and radiological findings. The frequency of drug related reactions will be tabulated by visit. Listings will be provided as well.

For laboratory assessments, the shifts relative to CTCAE criteria for laboratory abnormalities will be provided. The laboratory measures will also be compared with their corresponding normal ranges, and the incidence of abnormally high and abnormally low laboratory values will be summarized.

19. Subgroup Analysis

Primary and secondary endpoints which demonstrate statistical significance will also be analyzed by subgroups including sex (male/female), race, ethnicity, and disease severity (according to the definitions of the electronic Case Report Forms [eCRFs]). A subgroup analysis by grouped coded terms of concomitant medications may also be provided.

20. Change from Analysis Specified in the Corresponding Protocol Version

Not applicable.

REFERENCES

Mantel, N., and Fleiss, J. L. (1980). “Minimum Expected Cell Size Requirements for the Mantel-Haenszel One-Degree-of-Freedom Chi-Square Test and a Related Rapid Procedure.” *American Journal of Epidemiology* 112:129–134.

Miettinen, O. S., and Nurminen, M. M. (1985). “Comparative Analysis of Two Rates.” *Statistics in Medicine* 4:213–226.

Santner, T. J., and Snell, M. K. (1980). “Small-Sample Confidence Intervals for $p_1 - p_2$ and p_1/p_2 in 2×2 Contingency Tables.” *Journal of the American Statistical Association* 75:386–394.