NCT04707664

A Randomized Phase 2b Trial Evaluating Clinical Outcomes of Inhaled Sargramostim in High-Risk Patients with Mild-Moderate COVID-19

Sargramostim use in COVID-19 to Recover Patient Health (SCOPE)

Protocol Number: PTX-001-003

IND Number: 149323

Protocol Version: 4.0
Protocol Version Dated: 23 August 2021

Sponsor Name:

Partner Therapeutics, Inc.

19 Muzzey Street, Lexington, MA 02421

Statistical Analysis Plan: Version 1.0

SAP Version Date: 13 October 2021

STATISTICAL ANALYSIS PLAN APPROVAL PAGE

Document Information	
Version	1.0
Document Date	13 October 2021
Prepared for	Partner Therapeutics, Inc.
Prepared by	

Sponsor Approver details	
Name	
Job Role	Consultant – Regulatory, Biostatistics & Data Management
Company	Partner Therapeutics, Inc.
Signature	
Date of signature	
Name	
Job Role	Executive Director Medical Immunology Oncology
Company	Partner Therapeutics, Inc.
Signature	
Date of signature	

Other Approver details	
Name	
Job Role	Director of Data Sciences and Biostatistics
Company	FHI 360
Signature	
Date of signature	

TABLE OF CONTENTS

LIST OF TABLES

STUDY SUMMARY

Title:	A Randomized Phase 2b Trial Evaluating Clinical Outcomes of Inhaled Sargramostim in High-Risk Patients with Mild-Moderate COVID-19
Design:	Phase 2b, multicenter, placebo-controlled, double-blind, randomized study
Investigational Drug:	250 mcg inhaled sargramostim
Reference Treatment:	Equivalent volume of inhaled placebo diluent
Population:	Approximately 500 adult patients who are symptomatic with mild or moderate COVID-19 who are at high risk for progression to more severe disease. Enrollment of patients who have completed a COVID-19 vaccination regimen or participated in a COVID-19 vaccine clinical trial will be capped at approximately 100 patients.
Study Duration:	After randomization, each patient is expected to stay in the study for approximately 60 days. The study includes up to 5 days for screening, 5 days of treatment and follow-up through Day 60.Refer to Protocol Section 11 for the Schedule of Events.
Primary Objective:	To evaluate if inhaled sargramostim can prevent progression to more severe disease in an outpatient setting in symptomatic patients with mild or moderate COVID-19 who are at higher risk for progression to more severe disease
Primary Endpoint:	Proportion of patients who experience any emergency room visit, hospitalization, or death by Day 28
Secondary Objectives:	1. To evaluate the effect of inhaled sargramostim on clinical progression of COVID-19
	2. To evaluate the safety of inhaled sargramostim in patients with COVID-19
Secondary Endpoints:	1a. Proportion of patients with any progression of disease as determined by a ≥ 2-point increase from baseline in the NIAID ordinal scale up to Day 28, and Day 60
	1b. Time to progression of disease as determined by a ≥ 2 -point increase in the NIAID ordinal scale up to Day 28, and Day 60
	1c. Change from baseline in overall symptom score and individual symptom scores, as measured by the Symptom Score Questionnaire up to Day 28
	2. Adverse events (AEs) up to Day 60
Exploratory Objective:	To explore the effect of inhaled sargramostim on biological responses to COVID-19
Exploratory Endpoints:	Proportion of patients requiring supplemental oxygen up to Day 28 The effect of sargramostim on SARS-CoV-2 viral load in samples from nasopharyngeal swabs up to Day 14 (biomarker cohort only) Proportion of patients with generation of SARS-CoV-2 antibody up to Day 28 (biomarker cohort only) Proportion of patients with markers of cellular immunity, via immunophenotyping up to Day 14 (biomarker cohort only) Levels of ferritin, c-reactive protein (CRP), d-Dimer up to Day 28 Immunogenicity to sargramostim up to Day 28 (biomarker cohort only) Presence of absence of COVID-19 symptoms at Day 60

LIST OF ABBREVIATION AND SPECIALIST TERM

Abbreviation	Description
AE	Adverse Event
CI	Confidence Interval
CoV	Coronaviruses
COVID-19	Coronavirus-2019
CRP	C-reactive Protein
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events (from the National Cancer Institute)
DM	Data Management
DSMB	Data and Safety Monitoring Board
eCRF	Electronic Case Report Form
EDC	Electronic Data Capture
ER	Emergency Room
FAS	Full Analysis Set
ICU	Intensive Care Unit
IWRS	Interactive Web-based randomization System
ITT	Intent-To-Treat
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
NCI	National Cancer Institute
NIAID	National Institute of Allergy and Infectious Diseases
PT	Preferred Term
RS	Randomization Statistician
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
US FDA	United States Food and Drug Administration
WI	Work Instruction

1. INTRODUCTION

This statistical analysis plan (SAP) expands *Section 14 (Statistical Plan)* of the study protocol "A randomized Phase 2b trial evaluating clinical outcomes of inhaled sargramostim in high-risk patients with mild-moderate COVID-19". This SAP is the first version. Final SAP approval by the Sponsor will occur before data unblinding. Any changes to the planned analyses will be documented in the Clinical Study Report (CSR).

The scope of analysis provided in this SAP includes plans for final analysis after the last patient completes the Day 60 visit or an early termination visit. This SAP also includes a limited description of the planned interim analysis for Data and Safety Monitoring Board (DSMB) purposes after approximately the first 100 patients enrolled complete 28 days of the study or discontinue earlier from the study (see the DSMB Charter for more details).

2. STUDY OBJECTIVES

2.1. Primary Objective

To evaluate if inhaled sargramostim can prevent progression to more severe disease in an outpatient setting in symptomatic patients with mild or moderate COVID-19 who are at higher risk for progression to more severe disease.

2.2. Secondary Objectives

1. To evaluate the effect of inhaled sargramostim on clinical progression of COVID-19;
2. To evaluate the safety of inhaled sargramostim in patients with COVID-19;

2.3. Exploratory Objectives

To explore the effect of inhaled sargramostim on biological responses to COVID-19.

3. STUDY DESIGN AND PROCEDURES

This Phase 2b study will randomize approximately 500 patients who are symptomatic with mild or moderate COVID-19 who are at high risk for progression to more severe disease. Patients will be randomized in a 1:1 ratio to 250 mcg of inhaled sargramostim plus SOC or equivalent volume of placebo diluent plus SOC. Randomization will be stratified by geographic region (US, outside the US), and by completion of a COVID-19 vaccination regimen or participation in a COVID-19 vaccine clinical trial (yes, no). Enrollment of patients who have completed a COVID-19 vaccination regimen or participated in a COVID-19 vaccine clinical trial will be capped at approximately 100 patients.

Treatment will be administered once daily for 5 days and delivered via a vibrating mesh nebulizer. Additional blood samples will be collected for biomarker assays for approximately 100 randomized patients (approximately 50 from each treatment group). Patients will be followed for up to 60 days after start of treatment.

Refer to protocol sections 9 and 11 for a complete description of study visits and procedures along with a schedule of events.

4. GENERAL ANALYTIC CONSIDERATIONS

4.1. Study Size and Power

A total of approximately 500 patients will be randomized using a 1:1 assignment ratio. The sample size is selected to achieve at least 80% power to detect a difference at the 0.05 level of significance (2-sided) in the proportion of patients who experience any emergency room visit or hospitalization, or death over 28 days.

This sample size is based on an assumed 3% event rate for the sargramostim arm and a 10% rate for the placebo arm (risk ratio of 0.30) in unvaccinated patients (similar to that observed for bamlanivimab BLAZE-1 study [Chen, 2021; Gottlieb, 2021; US FDA, 2021]). As discussed in protocol Section 4.6, the rate of any emergency room visits or hospitalization, or death over 28 days is expected to be very low, or near zero for vaccinated patients.

It is also assumed that the proportion of vaccinated patients will be approximately 20-25% of the study population, with enrollment of patients who have completed a COVID-19 vaccination regimen or participated in a COVID-19 vaccine clinical trial capped at approximately 100 patients; no events are expected to occur in these patients for either treatment arm. Therefore, the overall event rate (for unvaccinated and vaccinated patients combined) for the primary endpoint is expected to be 2.4% for sargramostim and 8.0% for placebo (risk ratio of 0.30) and maintains at least 80% power.

Note that completion of a COVID-19 vaccination regimen or participation in a COVID-19 vaccine clinical trial (yes, no) is one of the randomization stratification factors (protocol Section 6.2).

4.1.1. Monitoring for potential impact of approved/authorized therapies on event rates

The Sponsor (or designee) will monitor (in a blinded manner and pooled across both treatment arms) the usage rate of authorized/approved therapies (such as monoclonal antibodies) for patients with mild or moderate COVID-19 in an outpatient setting following approval of protocol version 4.0. Given the observed efficacy that supports their regulatory authorization/approval, these therapies could impact this study's primary outcome measure. That is, patients who receive these therapies would be expected to have a very low or near zero rate of ER visit, hospitalization, or death.

If there is a notable rate of these therapies, a re-evaluation of sample size and power may be required. There will not be any unblinding of the rate of authorized/approved therapies (such as monoclonal antibodies), nor any unblinding of the primary endpoint (other than within the scope of the DSMB, see DSMB Charter). Therefore, no adjustment to Type 1 error is required. Any decisions around re-evaluation of sample size and power will be described in internal documentation and the CSR.

4.2. Data Source

Data management (DM) plans are detailed in a separate study-specific document. Briefly, clinical trial data from source documentation will be entered by the site into CRFs via a 21CFR Part 11 compliant electronic data collection system (EDC). AEs and concomitant medications will be coded according to the most current versions of WHODrug and MedDRA dictionaries. Patients will enter patient reported outcome survey data directly into the REDCap survey

database. A paper diary will be provided as a back-up data collection source for use in cases when the patient cannot utilize the REDCap system, either temporarily or permanently. In addition, laboratory and biomarker data will be transferred from the appropriate vendors, and compiled with the above data sources as part of the study database.

FHI 360 DM will create SAS datasets (version 9.4 or higher) containing study data for use by biostatisticians conducting interim and final analyses.

4.3. Test Size, Confidence Levels, and Precision

Unless otherwise noted, reported confidence intervals will all be computed at the 95% confidence level; p-values will be assessed at the two-sided 0.05 significance level. Unless otherwise specified below, continuous variables will be summarized using number of patients, mean, standard deviation, median, minimum, maximum, and 25^{th} and 75^{th} percentiles as appropriate. Categorical variables will be summarized using number and percentage of patients in each category. The number of patients with missing responses (or variables) will be tabulated.

As this is a Phase 2b study, no adjustments for multiple comparisons will be made.

Unless otherwise specified, the following guidelines will be followed when displaying statistics:

- measures of central tendency and percentiles will be displayed with one decimal place more than the raw (collected) data,
- measures of variability will be displayed with two decimal places more than the raw (collected) data,
- percentages will be displayed with one decimal place, and
- p-values will be displayed to 4 decimal places and any less than 0.0001 will be displayed as "<0.0001".

4.4. Strategy for Intercurrent Events

The treatment policy strategy will be employed for efficacy analyses. Analyses will consider all data, regardless of any intercurrent events, such as early discontinuation of therapy, any protocol deviations, or any additional treatments for COVID-19.

4.5. Missing Data

Rules for imputation or adjudication are described in detail in each endpoint definition below. Imputation rules for any partial dates are covered in Section 4.7. Additional rules for dealing with missing data may be determined during blinded data review and documented prior to unblinding.

4.6. Alternative Methods

Assumptions for each method outlined for each endpoint will be formally assessed, and if any underlying assumptions or distributions are not satisfied then alternative statistical methods will be employed. Any alternative methods used will be described in the CSR.

4.7. Derivations and Data Handling Conventions

Baseline will be defined as the last assessment prior to the first dose date/time (or randomization if patient was not treated).

Study Day will be derived as (assessment date – date of first dose) + 1.

Event durations in days will be derived as (end date – start date) + 1.

Partial dates will be imputed for derivation of time durations (e.g., hospitalizations, concomitant medications/procedures, adverse events, etc.), as follows:

- For start dates missing the day component, the first of the month will be used.
- For end dates missing the day component, the last of the month will be used.
- Any start dates missing the month component will not be imputed and durations will be listed as unknown unless the study team has evidence to provide a more reasonable date.
- Any end dates missing the month component will be imputed as the latest of study disposition date and death date (if applicable).
- Other collected study data may be used to inform imputation of missing date components.

5. RANDOMIZATION, ALLOCATION CONCEALMENT AND BLINDING

5.1. Randomization

Patients will be stratified by geographic region (US, outside the US), and by completion of a COVID-19 vaccination regimen or participation in a COVID-19 vaccine clinical trial (yes, no). The randomization sequence will be developed by a qualified FHI 360 Randomization Statistician (RS) who is not otherwise involved in the study using a validated program written in SAS®.

5.2. Allocation Concealment and Blinding

The randomization statistician will prepare an electronic randomization list containing treatment assignments following the randomization sequence to be uploaded to the Interactive Web Response System (IWRS) operated by the vendor, Clinical One. After patient eligibility has been determined, site staff will perform online randomization through Clinical One IWRS. Site staff will remain blinded and will be trained in proper randomization and unblinding procedures. As the study is a double-blind study, neither patients, nor investigators, nor the Sponsor/CRO study team will be aware of treatment assignments prior to the final data base lock at the conclusion of the study.

5.3. Blinded Data Review

Prior to scheduled study unblinding, the data may be reviewed in a blinded fashion per the Sponsor's or medical and safety monitoring stakeholder's request. Based on analysis needs, the lead statistician or data manager will develop specific outputs for this review and will document the Sponsor or stakeholders' decisions prior to unblinding. These reviews may include (but not necessarily be limited to) protocol deviations, violations of inclusion/exclusion criteria, and/or data points. Protocol deviations will include those documented by the study team as well as those that can be detected directly from the data base. Decisions on inclusion of patients into each analysis population, missing data, and final imputation of endpoints will be made during this blinded review. Blinded data review may also include clinical evaluation of safety data and to issue clinical queries to the site, if needed. Any decisions affecting analysis will be described in internal documentation and the CSR.

5.4. Scheduled End of Study Unblinding

After the study completes enrollment and patient follow-up, data entry, discrepancy resolution, blinded results reviews, and database lock, then the process for full study unblinding will start. Formal request to unblind should come from the Sponsor via a signed "Request to Reveal the Randomization Schedule" FHI 360 form. Once the Sponsor indicates that all requirements for full unblinding have been met, the lead biostatistician will initiate completion and sign-off of this form and document the unblinding process.

6. ANALYSIS POPULATIONS/SETS

The following analysis populations will be used when performing statistical analyses or summaries:

- Screened Population includes all patients who are screened.
- Full Analysis Set (FAS) includes all randomized patients and will be analyzed according to treatment as randomized. This set is analogous to an Intent-to-Treat Population. This population will be used as the primary analysis population for baseline and efficacy analyses.
- Modified Intent-to-Treat (mITT) Population includes all randomized patients who meet eligibility criteria and receive at least one dose of any study treatment. Patients will be excluded from the mITT Population if they fail to meet eligibility criteria or if they do not receive any study treatment. This population will be used for sensitivity analyses and patients will be analyzed according to their randomized treatment.
- Safety Population/Set includes all randomized patients who receive at least one dose of any study treatment. Safety analyses will be performed according to the treatment received on Day 1.
- Biomarker Analysis Population includes all randomized patients who consent to provide additional samples for biomarker assessment and who receive at least one dose of any study treatment. Biomarker analyses will be performed according to the treatment received on Day 1.

A summary of the planned study analysis by patient population is displayed in Table 1.

Table 1. Summary of Planned Analyses, by Population

Analysis	Screened	FAS	mITT	Safety	Biomarker
Participant Disposition	√	√			
Demographics and Baseline		√	√	√	√
Prior and Concomitant Medications, Procedures, and Surgeries		√			
Medical History		√			
Intercurrent Events		√			
Primary Outcome Measure (Efficacy)		√	√ (sensitivity)		
Secondary Outcome Measures (Efficacy)		√	If warranted (see Section 12.2)		
Other Efficacy Data (see Section 9.3)		√			
Secondary Outcome Measures (Safety)				√	
Exploratory Outcome Measures (Efficacy, non-biomarker)		√			
Exploratory Outcome Measures (Safety, non-biomarker)				√	
Exploratory Outcome Measures (Biomarker, Biomarker cohort only)					√

7. DATA AND SAFETY MONITORING BOARD MEETING SCOPE AND TIME POINT

An independent data and safety monitoring board (DSMB) will be convened for study oversight. The DSMB charter is a separate document. The DSMB will review unblinded data after approximately the first 100 patients have completed 28 days in the study or discontinued from the study. They will review available safety and efficacy data for a benefit/risk evaluation, and data will be accepted as is with the understanding that it may not have gone through a full data cleaning process. No formal interim efficacy analysis will be performed for the purpose of stopping the study. Rather, if the DSMB perceives an unfavorable benefit/risk to the patients in the study or overall lack of efficacy, they may recommend that the Sponsor take appropriate actions, including enrollment pause, modifying or stopping the study, and/or requesting an additional DSMB data review and meeting. Refer to the DSMB charter for more information.

The DSMB charter contains a list of suggested data outputs to be included in the DSMB Report. Patient disposition will be summarized using the Screened Population. All other summaries will be provided using the Safety Population. Continuous data will be summarized by number of patients, mean, standard deviation, median, minimum, maximum, and 25^{th} and 75^{th} percentiles where appropriate. Categorical data will be summarized by number and percentage of patients with available data. Table and listing shells for the DSMB meeting will be reviewed and approved by the Sponsor in advance of the DSMB meeting.

All data presented to the Sponsor will be blinded and aggregated across all treatment groups; that is, the Sponsor and study team will not be unblinded. The DSMB members will receive an unblinded version of the report that will be produced by an unblinded statistician that is not otherwise involved in the study.

8. DESCRIPTION OF PARTICIPANT POPULATION

8.1. Disposition

Number and percentage of patients screened, screened but not enrolled, randomized, and treated will be tabulated based on the Screened Population. Numbers and percentages of patients who qualify for each analysis population will also be tabulated, along with any reasons for exclusions. Patients' final disposition will be tabulated using the FAS and by treatment arm, including frequency and percentage of patients who complete the study, are lost to follow-up, or terminate early along (with reason for discontinuation), as well as key visit completion. Patient level listings will also be provided for disposition data for the Screened Population. If any patients experience treatment cross-over between doses (i.e., they receive a different dose at a later treatment than they received at their first treatment), a listing will be provided for such patients in the Safety Population which describes the actual treatment received on each day.

There may be instances of discordance in the randomization stratum values entered during the randomization process, and what is verified by the study site (eCRF). If this occurs, the statistical analyses outlined in this SAP will be performed using the values of the strata that are site verified on the eCRF. A concordance table will evaluate agreement/disagreement between values of the randomization strata between what was used during randomization, and what is verified by the site; this tabulation will show both treatment arms as well as the overall study population (using the FAS).

8.2. Protocol Deviations

Protocol deviations management is documented in a separate Protocol Deviation Plan. Major protocol deviations will be tabulated and summarized for the FAS. Protocol deviations will be listed including classification as major or minor along with patient ID, treatment arm, age, and sex for the Screened Population.

8.3. Demographics and Baseline Characteristics

Baseline demographic information, baseline height, weight, BMI, and baseline disease characteristics, will be summarized by treatment group for the FAS, mITT, Safety, and Biomarker Analysis Populations. Baseline disease characteristics will include SARS-CoV-2 variant, oxygen saturation on room air, andCOVID-19 symptoms. The number of patients with missing data for a given variable will be tabulated. Patient level listings will also be provided for the FAS with an indication of randomization, treatment, and eligibility status.

8.4. Prior and Concomitant Medications, Procedures and Surgeries

Prior medications are defined as medications started within 4 weeks of the screening visit up to the start of study drug, and ended prior to Study Day 1. Concomitant medications are defined as medications administered during the patient's participation in the study from informed consent date until Day 60; that is, the time between start and/or stop date of the concomitant medication includes any time from screening to Study Day 60. The same concept applies to procedures and surgeries to determine which are prior or concomitant.

Prior and concomitant medications will be summarized by treatment group for the FAS. As selected COVID-19 standard of care may begin prior to patient's informed consent, all prior and concomitant medications will be included.

Frequency summaries of coded prior and concomitant medications will be provided for each drug class and preferred base term by treatment group separately for selected COVID-19 standard of care medications and all non-COVID-19 medications. In order to ensure consistency across sites, all recorded medications will be reviewed in a blinded fashion by the study team to determine which are considered selected standard of care for COVID-19. This determination will be provided to the data management/statistical team.

At each level of summarization, each patient will be counted once if they reported one or more medications. A prior and concomitant medication listing, including randomization and treatment, and a flag for selected COVID-19 standard of care medications, reason for use, duration, frequency, route, and dosage will be provided.

Prior and concomitant procedures and surgeries will be summarized by treatment group for the FAS.

All prior and concomitant procedures and surgeries will be listed by patient for the FAS with an indication of randomization, treatment, and procedure/surgery indication.

8.5. Medical History

Baseline medical history data (including preexisting conditions) will be summarized using frequency tabulations by the Medical Dictionary for Regulatory Activities (MedDRA) system organ class (SOC) and preferred term (PT) for the FAS. Each patient will only be counted once per SOC and once per PT. Patient level listings will also be provided for the FAS with any missing or partial dates presented as shown in the study data.

8.6. Intercurrent Events

Intercurrent events such as early discontinuation of treatment, any treatment crossover, any major protocol deviations (see Section 8.2), and any other approved or authorized treatments to prevent progression of COVID-19 disease severity or to treat COVID-19 (such as monoclonal antibodies, antivirals such as remdesivir, etc.) will be tabulated by treatment group for the FAS. Patient level listings will also be provided for the FAS.

9. ASSESSMENT OF EFFICACY

Efficacy endpoints will include need for emergency room visit or hospitalization, death, the NIAID ordinal scale, and evaluation of symptoms using symptom score and patient global impression questionnaires (Section 2).

9.1. Analysis of Primary Outcome Measure

Primary efficacy analysis will be done using the FAS. The primary endpoint is the proportion of patients who experience any emergency room (ER) visit, hospitalization, or death by Day 28. Any occurrence of any of the 3 types of events as recorded on the Hospitalization (with healthcare type marked as emergency room, hospitalization, or intensive care unit (ICU)) or Death eCRFs will be considered an event, and each patient will only contribute one event to the analysis. Duration of ER visit and hospitalization stay (including any ICU stay) based on duration < 24hrs or ≥ 24 hrs will be tabulated separately. Stays in extended care facilities will also be recorded on the eCRFs but these will not count towards the primary endpoint. It is assumed that a stay in such a facility would follow a hospitalization and that hospitalization would count toward the endpoint. Any responses of "other" to the healthcare type will be reviewed and categorized by the study team prior to unblinding. Any patient in the FAS that is lost to follow-up prior to Day 28, such that determination of ER visit, hospitalization or death cannot be made, will be considered as no event in the primary analysis.

Statistical hypothesis testing for the primary endpoint will be conducted using a logistic regression model with a Firth penalized likelihood (Firth, 1993) at the two-sided 0.05 significance level. The model will include treatment as the independent variable. Event rates within strata will be evaluated to determine whether it is appropriate to conduct a sensitivity analysis that includes stratification by the randomization stratification factors for geographic region (US vs outside US) and completion of a COVID-19 vaccination regimen or participation in a COVID-19 vaccine clinical trial (yes, no). This analysis may include an exact conditional logistic regression with stratification by appropriate variables or other similar methodology. Based on the observed event rates within strata, the decision whether to conduct this stratified sensitivity analysis and the methodology to be used will be made prior to any unblinding of the study team and documented in the final study report.

The primary estimand of treatment effect will be the risk difference between treatment arms, corresponding 95% confidence interval (CI), and p-value based on a Wald test. The confidence interval for the risk difference will be obtained using the delta method. Additionally, the odds ratio, corresponding 95% CI, and p-value will be generated.

The proportion of patients who have any ER visit, hospitalization, or death by Days 7, 14 and 28, and at any time during the study (e.g., to Day 60) will be tabulated. A Chi-square or Fisher's exact test will compare treatment arms at each timepoint of interest. Number and percentage of patients who experience each type of healthcare as listed on the Hospitalization eCRF will also by tabulated by treatment arm.

Individual data listings of ER visits, hospitalizations, or deaths will be provided with an indication of whether or not the event occurred by Day 28 and whether or not the event was COVID-19-related, along with treatment arm, randomization strata, age, race, and sex for each patient.

Two sensitivity analyses will be conducted mirroring primary endpoint analysis methods as specified above.

1. The first sensitivity analysis will define the endpoint as any COVID-19-related emergency room visits, COVID-19-related hospitalizations, or deaths (any cause).
2. The second sensitivity analysis will define the endpoint the same way as the primary analysis, but will be conducted using the mITT Population.

A third sensitivity analysis will be conducted if the primary endpoint is missing for 5% or more of patients; i.e., missing the primary endpoint, in this case, refers to patients who were not followed through Day 28, and were not known to have an ER visit, hospitalization, or death. This analysis will use a multiple imputation approach (such as Marcov Chain Monte Carlo method, or FCS method with logistic regression) to impute missing values for patients who are missing the primary endpoint due to loss to follow-up or discontinuation prior to Day 28 (Little 2019). Covariates to consider when imputing the missing data will at a minimum include treatment arm, and those listed in Section 12.1.

9.2. Analysis of Secondary Outcome Measures

Secondary efficacy analyses will be performed using the FAS unless otherwise specified. Individual data listings of NIAID ordinal scale results, symptom score questions and overall symptom score will be provided for each available day of record along with treatment arm, randomization strata, age, race, and sex for each patient.

In the event of missing NIAID ordinal assessments, any available hospitalization records, supplemental oxygen use information, or death information, as well as patient reported outcomes, will be used to impute the NIAID ordinal score if possible, particularly for ordinal scores by Days 7, 14, 28 and 60. This imputation will be conducted in a blinded fashion by the medical monitor and results will be provided to the data management/statistical team. If there is a missing assessment that is unable to be imputed using the data described above, but the missing assessment occurs between two non-missing assessments, then the higher non-missing score will be imputed for the missing score. Imputed scores will be used in the analyses below (unless specified otherwise, i.e., for sensitivity analyses) and any missing scores that are unable to be imputed will be ignored. If less than 5% of expected NIAID ordinal assessments are missing and unable to be imputed, then any impact on analysis will be considered negligible and no further action will be taken. If 5% or more are missing and unable to be imputed, then the study team will consider imputing an appropriate “high score” that is reasonable based on other responses. This “high score” will be chosen and documented prior to unblinding.

9.2.1. Proportion of patients with any progression of disease as determined by a ≥ 2-point increase in the NIAID ordinal scale up to Day 28 and Day 60

A secondary endpoint is the proportion of patients with progression of disease as determined by a ≥ 2-point increase in the NIAID ordinal scale up to Days 28 and 60 (e.g., a change from 2 to 4, or a change from 2 to 5). The analysis will be performed for each timepoint using a logistic regression model with a Firth penalized likelihood (Firth, 1993) at the two-sided 0.05 significance level. The model will include treatment as the independent variable. Event rates within strata will be evaluated to determine whether it is appropriate to conduct a sensitivity analysis which includes stratification by the randomization stratification factors for geographic

region (US vs outside US) and completion of a COVID-19 vaccination regimen or participation in a COVID-19 vaccine clinical trial (yes, no).

The primary estimand of treatment effect will be the risk difference between treatment arms, corresponding 95% confidence interval (CI), and p-value based on a Wald test. The confidence interval for the risk difference will be obtained using the delta method. Additionally, the odds ratio, corresponding 95% CI, and p-value will be generated.

The proportion of patients with any progression of disease as defined above by Days 7, 14, 28 and 60 will be tabulated. A Chi-square or Fisher's exact test will compare treatment arms at each timepoint of interest.

A sensitivity analysis will be conducted mirroring analysis methods described above but excluding any imputed scores. A repeated measures modelling approach may also be considered to utilize the available longitudinal data.

9.2.2. Time to progression of disease as determined by a ≥ 2-point increase in the NIAID ordinal scale up to Day 28 and Day 60

Time to progression of disease will also be assessed. Time to progression is defined as time from randomization to first occurrence of a ≥ 2-point increase in the NIAID ordinal scale up to Day 60. Patients who do not have a ≥ 2-point increase in the NIAID ordinal scale by Day 60 will be censored at the date of their last non-missing NIAID assessment up to Day 60. Kaplan-Meier methods will be used for the first time to event analysis. A log rank test stratified by the 2 randomization stratification factors, as appropriate, will be used to compare the 2 treatment arms. Estimates for probability of progression at Days 7, 14, 28, and 60 days, as well as median time to progression, will be derived from Kaplan Meier analysis.

Cox's proportional hazards model stratified by the 2 randomization stratification factors, as appropriate, will also be used to derive a hazard ratio and 95% CI assessing treatment effect, along with comparisons of the timepoint estimates at Days 28 and 60, if warranted. The proportional hazards assumption necessary for both the log rank test and Cox's proportional hazards model will be assessed. Any alternative methods used will be described in the CSR.

Two sensitivity analyses will be conducted mirroring analysis methods described above:

1. Excluding any imputed scores.
2. Imputing missing data when there is no collected data, as follows:
    a. If there are missing scores prior to first occurrence of a ≥ 2-point increase in the NIAID ordinal scale, then it will be assumed for purposes of this sensitivity analysis that the patient's score would have been equal to whatever score was recorded directly after the missing score

9.2.3. Change from baseline in overall symptom score, and individual symptom scores, as measured by the symptom score questionnaire up to Day 28

Change from baseline in the overall symptom score, and individual symptom scores from baseline to Day 28 will be analyzed using a mixed model repeated measures analysis using a missing at random assumption for handling missing data. The model will contain baseline symptom score, treatment arm, day, and treatment-by-day interaction as fixed effects, along with strata for the two randomization factors, as appropriate. The primary estimand of treatment effect

will be the treatment differences and corresponding 95% CIs, with p-values. Estimates and corresponding 95% CIs at days 7 and 14 will also be generated. Numbers and percentages of patients with each score at baseline, and days 7, 14, and 28 will also be summarized for the overall score and for each symptom by treatment arm. All available data will be used in the analysis; no imputation will be done for missing data.

The overall symptom score will be calculated as a sum of individual scores from 14 questions on the questionnaire. Each response will be scored according to the FDA's Guidance on "Assessing COVID-19-Related Symptoms in Outpatient Adult and Adolescent Subjects in Clinical Trials of Drugs and Biological Products for COVID-19 Prevention or Treatment" (US FDA, 2020). The scores correspond to the responses:

Questions 1-10: None = 0, Mild = 1, Moderate = 2, Severe = 3;

Questions 11-12: Not at all = 0, 1-2 times = 1, 3-4 times = 2, 5 or more times = 3;

Questions 13-14: Same as usual = 0, Less than usual = 1, No sense = 2.

A sensitivity analysis will be conducted mirroring the analysis methods described above but using the following assumptions:

- Completely exclude from analysis any patients who are missing more than 20% of their required symptom score questionnaires, unless the days where they are missing questionnaires fall after the patient experienced hospitalization, ER visit, or death.
- In patients with missing scores that account for 20% or less of their required symptom score questionnaires (or patients whose scores are missing due to prior hospitalization, ER visit, or death), handle as follows:
  - For any individual symptom scores that are missing because of the patient's prior death, impute the most severe possible answer.
  - If a patient has missing symptom scores during a period of hospitalization or ER visit, also impute the most severe possible answer for individual symptom scores.
  - For patients missing scores for any other reason, the higher of the two surrounding scores will be imputed for the missing individual symptom score (or scores).
  - The overall score will be calculated as the sum of individual symptom scores (reported or imputed) if possible.

Means and 95% CIs for the overall symptom score will be plotted over time from baseline through Day 28 by treatment arm (e.g., linear plot over time). Individual symptom scores will be summarized graphically using stacked bar charts at baseline and days 7, 14, and 28.

9.3. Analysis of Other Efficacy Data

These other efficacy data comprise efficacy outcomes of interest that are not strictly captured as primary, secondary, or exploratory outcomes measures of efficacy. This includes the patient global impression score. Note, the exploratory outcome measure of presence/absence of COVID-19 symptoms at Day 60 is included in Section 11.3.

These other efficacy data will be presented for the FAS.

Analysis of the patient global impression questionnaire will be presented by treatment group. This form was revised part way through the trial in protocol version 3 and went from one question to three. All available data will be analyzed and the responses in the final analysis will include the questions asked on the first version of the questionnaire as well as the second. Any participants who were enrolled prior to the time when their site implemented protocol version 3 will be summarized as missing for questions 1 and 3 on the second version of the questionnaire for the relevant timepoints and excluded from any time to event analysis for those questions. Answers to the single question on the first version of the questionnaire will be combined with answers to the second question on the second version of the questionnaire.

Median time to return to usual health, median time to return to usual activities, and median time to no COVID-19-related symptoms will be summarized for each treatment group using Kaplan-Meier methods. Responses on days 7, 14 and 28 will also be summarized using shift tables from baseline. The proportion of patients who return to usual health, return to usual activities, and with no COVID-19-related symptoms will be tabulated by Days 7, 14, 28, and 60; a Chi-square or Fisher's exact test will compare the treatment arms for each question at each timepoint.

If patients are missing the patient global impression questions at all timepoints, then for time to event analyses, they will be censored (not returned to usual health, activities and/or symptoms not resolved) at Day 60 or at date of study discontinuation, if earlier than Day 60.

10. ASSESSMENT OF SAFETY (SECONDARY OUTCOME MEASURES)

All safety analyses will be performed using the Safety Population. All hospitalizations, emergency room visits, and deaths will be considered in the efficacy rather than safety analyses for this study. All patients will be assessed regularly for potential adverse events (AEs) occurring from the time the patient provides informed consent until participation in the study has ended (up to 60 days after the first dose of study medication); see Protocol Section 10 for Adverse Event definitions for this study. All AEs reported in the study database will be included in summaries, regardless of whether they are determined to be associated with COVID-19.

Other safety analyses will summarize treatment administration and exposure, laboratory parameters, vital signs, and pulse oximetry (oxygen saturation).

No missing data will be imputed for safety analysis nor will formal hypothesis testing be performed.

Treatment-emergent AEs (TEAEs) will include AEs which newly emerge on or after the first study drug dosing or a preexisting medical condition judged by the Investigator to have worsened in severity or frequency or changed in character on or after the first study drug dosing.

A treatment-related AE is defined as an AE which is considered to have a possible, probable, or definite relationship to the study drug. A device-related AE is defined as an AE which is considered to have a possible, probable, or definite relationship to the nebulizer. AEs will be graded according to the National Cancer Institute's Common Terminology Criteria for Adverse Events (NCI CTCAE).

Patient level listings of AEs, serious AEs (SAEs), study treatment duration and exposure, laboratory parameters (including abnormal flags), vital signs, pulse oximetry and physical examination will be provided for all available assessments. Each listing will include patient ID, treatment arm, randomization strata, age, race, and sex. Any partial or missing dates will be reported as recorded in study data on patient listings.

Patient profiles may be generated for patients with SAEs, ER visits, hospitalizations, early discontinuation from treatment, and deaths. These patient profiles will include demographics and study treatment data, along with key laboratory data, concomitant medications, medical history, and adverse event data. Additional profiles may be generated for any identified suspected unexpected serious adverse reactions.

10.1. Summary of Study Treatment Duration and Exposure

All administered doses will be listed and summarized descriptively for each treatment day and treatment arm. Duration of study treatment for each patient will be summarized, with duration defined as the number of days on which the patient received treatment, up to a total of five. The number and percentage of patients per treatment arm that complete all 5 planned doses will be tabulated.

10.2. Analysis of Secondary Safety Endpoint

A secondary trial objective is to evaluate safety of inhaled sargramostim in patients with COVID-19 using adverse events (AEs) up to Day 60. Numbers and percentages of patients experiencing AEs will be presented by treatment arm.

An overview of the number of events and number and percentage of patients with AEs, SAEs, TEAEs, TEAEs related to study treatment or device, TEAEs related to COVID-19, serious TEAEs, fatal TEAEs, TEAEs that led to treatment discontinuation, and TEAEs by severity will be tabulated by treatment arm and overall.

TEAEs will be summarized as follows, by MedDRA system organ class (SOC) and preferred term (PT) and overall:

- All TEAEs
- All TEAEs by study period (dosing/post-dosing)
- All TEAEs by maximum severity (CTCAE toxicity grade)
- All treatment-related TEAEs by maximum severity (CTCAE toxicity grade)
- All device-related TEAEs by maximum severity (CTCAE toxicity grade)
- Serious TEAEs
- Treatment-related serious TEAEs
- Fatal TEAEs (CTCAE Grade 5 and/or an outcome of fatal)
- TEAEs leading to discontinuation of treatment

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

10.3. Analysis of Additional Safety Data

Safety laboratory parameters will include hematology and chemistry parameters collected at baseline, Day 14 and Day 28. Hematology parameters will include complete blood count with differential, hemoglobin, hematocrit, and absolute counts for white blood cells, platelets, neutrophils, lymphocytes, eosinophils, and monocytes. Chemistry parameters will include albumin, blood urea nitrogen, creatinine, alanine aminotransferase, aspartate aminotransferase, lactate dehydrogenase, alkaline phosphatase, direct and total bilirubin, glucose, total protein, sodium, potassium, chloride, bicarbonate, calcium, and uric acid. Additional safety laboratory parameters will include serum ferritin, CRP, and d-Dimer (planned analyses for these parameters are included in Section 11). Vital sign assessments will include temperature, blood pressure, pulse rate, and respiration rate. Pulse oximetry will also be summarized. Results for pulse oximetry may be influenced by altitude of clinical sites, so a subgroup analysis of patients enrolled at high altitude sites may be warranted (if there are sufficient numbers of patients).

Laboratory parameters, vital signs, and oxygen saturation (pulse oximetry) and change from baseline at each time point will be presented using summary statistics (n, mean, standard deviation [SD], median, interquartile range, minimum, and maximum values) by treatment group. Abnormal laboratory findings, will also be tabulated using percentages for below and above the normal range. No inferential testing will be performed to compare the differences between the treatment groups.

11. ANALYSIS OF EXPLORATORY OUTCOME MEASURES

The exploratory objective stated in the protocol is to explore effects of inhaled sargramostim on biological responses to COVID-19.

11.1. Exploratory Efficacy Outcome Measures

11.1.1. Proportion of patients requiring supplemental oxygen up to Day 28

Numbers and proportions of patients requiring any supplemental oxygen up to Day 28 will be summarized by treatment arm along with a summary by modality of oxygen supplementation support required for the FAS. For modality of oxygen supplementation required, each patient will be counted once per type required even if they required support in more than one instance. Number of days on oxygen support will also be summarized, for the FAS, with the number of days calculated using the difference between the start and end date-times recorded on the eCRF. Patient level listings of supplemental oxygen use will be provided for the FAS including patient ID, treatment arm, randomization strata, age, race, and sex. Any partial or missing dates will be reported as recorded in the study data on patient listings.

11.1.2. Presence or absence of COVID-19 symptoms at Day 60

The presence or absence of COVID-19 symptoms at Day 60 will be assessed for the FAS. This questionnaire data will be tabulated by question (symptom), in terms of the number and proportion of patients who have specific symptoms reported in the 7 days prior. A summary of the number of patients who completed or missed the Day 60 visit will also be provided. This symptom assessment is expected to be completed for patients at study discontinuation, if earlier than Day 60. There will be no imputation of missing data for this analysis. Patient level listings of presence or absence of COVID-19 symptoms at Day 60 will be provided for the FAS including patient ID, treatment arm, randomization strata, age, race and sex. Any partial or missing dates will be reported as recorded in the study data on patient listings. The data may also be presented graphically such as using a stacked bar chart, etc. by question (symptom) and by treatment arm.

11.2. Exploratory Safety Outcome Measures: Levels of ferritin, c-reactive protein (CRP), d-Dimer up to Day 28

Levels of ferritin, CRP, and d-Dimer on Days 1, 5, 14, and 28 will be summarized by treatment arm along with change from baseline for the Safety Population. The number of patients with missing values will be tabulated. Patient level listings of levels of ferritin, CRP, and d-Dimer up to Day 28, will be provided for the Safety Population including patient ID, treatment arm, randomization strata, age, race, and sex. Any partial or missing dates will be reported as recorded in the study data on patient listings.

11.3. Exploratory Outcome Measures for the Biomarker Cohort Only

Four additional measures will be assessed only for the Biomarker Analysis Population. These measures are: the effect of sargramostim on SARS-CoV-2 viral load in samples collected from nasopharyngeal swabs up to Day 14; proportion of patients with generation of SARS-CoV-2 antibody up to Day 28; proportion of patients with markers of cellular immunity, via immunophenotyping, up to Day 14; and immunogenicity to sargramostim up to Day 28. The

planned analysis of biomarker data for the biomarker cohort will be detailed separately in a Biomarker SAP Addendum, and finalized prior to any unblinding of the study team.

12. EXPLORATORY ANALYSES

12.1. Subgroup Analysis and Adjustment for Covariates

All subgroup analyses and covariate adjustments will be treated as exploratory since the study was not powered for such analyses. These analyses will be conducted using the FAS, unless otherwise specified. The following baseline characteristics will be considered for subgroup analyses and covariate adjustment in statistical models. Additional baseline factors may be considered for these analyses and a decision on which factors to include will be made while the study team remains blinded.

- Age group (18-59, ≥ 60 years). For regulatory purposes, an additional analysis may be performed using age groups defined as 18-64, and ≥ 65 years. If there are enough patients ≥ 75 years or ≥ 85 years, these categories may be included.
- Number of pre-existing medical conditions which put patients at higher risk of progression to severe COVID-19 (0-1, 2 or more), based on Inclusion Criterion 3b. Note: age will be considered separately as denoted above. If there are enough patients for any specific pre-existing condition in Inclusion Criterion 3b, a subgroup analysis may be performed for this.
- Duration of any COVID-19 symptoms (0-2, 3-5 days)
- Baseline NIAID ordinal score (1, 2)
- Gender (male, female)
- Race (White, Black or African American, Asian, all others)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Body mass index (<18.5 kg/m2, 18.5 to <25 kg/m2, ≥25 to <30 kg/m2, ≥30 kg/m2)
- Baseline SARS-CoV-2 variant (groupings to be determined based on prevalence)
- Use of approved/authorized COVID-19 therapies prior to randomization (if there are sufficient number of patients)

In addition to the above baseline factors, the following 2 randomization stratification factors will be used for subgroup analyses and included in all covariate adjustment models if there are sufficient patients in each of the strata.

- Region (US, outside US)
- Completion of a COVID-19 vaccination regimen or participation in a COVID-19 vaccine clinical trial (yes, no).

Prior to performing analyses, blinded review of baseline characteristics and randomization stratification factors will be performed to evaluate whether there are enough patients in each category. If there are insufficient numbers of patients in each category, some categories may be combined. Conversely, if there are more patients than expected in each category, additional categories may be considered. Additionally, collinearity between factors will be assessed prior to beginning any covariate adjustment. Any missing baseline factors will be accounted for using a missingness indicator (Groenwold et al. 2012).

Subgroup analyses will be conducted for the primary endpoint, and if warranted, for other efficacy endpoints. If there is an indication of differential treatment effect for the primary endpoint, then the subgroup analyses may be performed on secondary endpoints. Subgroup by treatment interactions will be included to assess differential treatment effects. If a meaningful differential treatment effect is observed across efficacy endpoints, then safety analysis may be performed for the subgroup(s) of interest. Furthermore, as the COVID-19 science continues to evolve, additional subgroups of interest may be identified.

As with the subgroup analyses, covariate adjustment will be performed for the primary endpoint, and if warranted, for other efficacy endpoints. A logistic or Cox proportional hazards regression model (based on type of endpoint) as described in Section 9.1 or Section 9.2 and will occur in 2 stages:

1. First, each baseline factor will be assessed in a model including treatment arm and selected randomization stratification factors to determine if each baseline factor has a potential impact on the outcome of interest. Factors will be assessed based on a univariate p-value < 0.1.
2. Next, all baseline factors with a univariate p-value < 0.1 will be added into a multivariate model containing treatment and randomization stratification factors. A stepwise approach using a significance level of 0.05 to enter the model and a significance level of 0.1 to remain will be used to determine the final covariate-adjusted model for each outcome of interest.

12.2. Other Exploratory Analyses

In addition to the previously specified exploratory analyses, additional exploratory analyses may be performed to evaluate the robustness and sensitivity of the study results, including but not limited to the analysis populations, subgroup analyses, treatment interactions, adjusted or stratified analyses, and/or alternative statistical methods.

For example, if there are meaningful differences in the results for the primary endpoint based on the mITT population compared to the analyses based on the FAS, then analyses of the secondary endpoints may be performed for the mITT population.

13. CHANGE FROM ANALYSIS SPECIFIED IN THE CORRESPONDING PROTOCOL VERSION

Section 12.1 includes a small change from the protocol in the covariate adjustment process. Instead of assessing baseline factors at a significance level of 0.05, factors will be assessed for inclusion into the stage 2 modeling based on a significance level of 0.1. This change was made due to an acknowledgement that some factors that may not be significant individually at a 0.05 level might be significant in a multivariate model.

14. TABLES, LISTINGS AND FIGURES

Planned tables, listings and figures (TLFs) intended to capture interim and final analyses described in the SAP will be reviewed by the Sponsor before going into production. TLF shells are maintained separately from the approved version of the SAP.

14.1. Programs and Quality Control

Report production and review will be conducted in accordance with FHI 360 BIOS Work Instructions (WI) 03003 (Verification of Analyses and Reports) and 03006 (Preparation and Review of Statistical Reports). In brief, a primary statistician-programmer for a given output will carefully review the program and output, verifying that no error message is highlighted in the "LOG" file and that titles, footers, footnotes, text body, etc. are correct. A second statistician-programmer will independently validate the output by checking the results against separately created SAS programs and checking textual material. Prior to delivery of any statistical output to Sponsor, the lead statistician will review the statistical package for internal inconsistencies or any items where clarifying notes will be helpful to the reviewer. Any package of final TLFs will also be thoroughly reviewed by the Director of Biostatistics and Data Sciences or designee before it is distributed.

14.2. Programming Conventions and Formatting

Reporting conventions will adhere, when possible, to the International Conference on Harmonization Guidance document E3, "Structure and Content of Clinical Study Reports".

All TLFs will be in landscape format unless otherwise specified. All listings will include patient ID, treatment arm, randomization strata, age, race, and sex unless otherwise approved by the study team.

Each TLF will have at least two titles: the TLF number and description, and identification of the study population.

Footnotes will include the following information:

- Date of data extraction or lock (once applicable)
- Run date
- Study number and/or name
- Report name
- Name of program used to create the file
- Listing number for associated listing (for tables and figures)

15. REFERENCES

Chen P, Nirula A, Heller B, Gottlieb RL, Boscia J, Morris J, et al. SARS-CoV-2 Neutralizing Antibody LY-CoV555 in Outpatients with Covid-19. N Engl J Med. 2021;384(3):229-37.

Firth D. Bias reduction of maximum likelihood estimates. Biometrika. 1993;80(1):27-38.

Gottlieb RL, Nirula A, Chen P, Boscia J, Heller B, Morris J, et al. Effect of Bamlanivimab as Monotherapy or in Combination With Etesevimab on Viral Load in Patients With Mild to Moderate COVID-19: A Randomized Clinical Trial. JAMA. 2021;325(7):632-44.

Groenwold RHH, et al. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. CMAJ. 2012. DOI: 10.1503/cmaj.110977

Little RJA, Rubin, DB. (2019). Statistical Analysis with Missing Data, 3rd edition, New York: John Wiley.

U.S. Food and Drug Administration. Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research. (September 2020). *Assessing COVID-19-Related Symptoms in Outpatient Adult and Adolescent Subjects in Clinical Trials of Drug and Biological Products for COVID-19 Prevention or Treatment. Guidance for Industry.* Retrieved from https://www.fda.gov/media/142143/download

U.S. Food and Drug Administration. Eli Lilly and Company [Fact Sheet For Health Care Providers Emergency Use Authorization (EUA) Of Bamlanivimab]. Indianapolis, IN 2021.