STATISTICAL ANALYSIS PLAN

VERSION: 5

DATE OF PLAN:

24-JUN-2024

BASED ON:

AT-03A-017 Protocol Version 2.0, 15-Jun-2023

STUDY DRUG:

Bemnifosbuvir Hemisulfate (BEM; AT-527)

PROTOCOL NUMBER:

AT-03A-017

STUDY TITLE:

A Phase 3 Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of Bemnifosbuvir in High-Risk Outpatients with COVID-19

SPONSOR:

Atea Pharmaceuticals, Inc.

225 Franklin Street

Suite 2100

Boston, MA 02110 USA

+1 (857) 284-8891

This study is being conducted in compliance with good clinical practice, including the archiving of essential documents.

SIGNATURE PAGE

This document has been prepared and reviewed by:

[redacted]

This document has been reviewed and accepted by:

[redacted]

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

1. LIST OF ABBREVIATIONS

Table 1: List of Abbreviations

Abbreviation	Term
AE	Adverse Event
ALT	Alanine Aminotransferase
AST	Aspartate aminotransferase
AUC	Area under the curve
BEM	Bemnifosbuvir
BID	Two times a day
BMI	Body mass index
CMH	Cochran-Mantel-Haenszel
CRO	Contract Research Organization
CSR	Clinical Study Report
DAIDS	Division of AIDS
DDI	Drug-drug interaction
DSMB	Data and Safety Monitoring Board
EFU	Extended follow-up
EOT	End-of-Treatment
EUA	Emergency use authorization
FDA	United States Food and Drug Administration
FDD	First dosing date of study drug
I/E	Inclusion/exclusion
ITT	Intent-to-Treat Population
LOD	Limit of Detection
LLoQ	Lower Limit of Quantification
mAb	Monoclonal antibody
MedDRA	Medical Dictionary for Regulatory Activities Terminology
MedDRASOC	System organ class
mITT	Modified intent-to-treat
mITT-C	Modified intent-to-treat for combination antiviral stratum
mITT-S	Modified intent-to-treat for supportive care stratum
NP	Nasopharyngeal

Abbreviation	Term
PK	Pharmacokinetics
PT	Preferred term
RNA	Ribonucleic acid
PP	Per-Protocol Population
ROW	Rest of world
RT-qPCR	Reverse transcription quantitative real-time polymerase chain reaction
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
SD	Standard deviation
SOC	Standard of care
SOP	Standard operating procedure
ULN	Upper Limit of Normal
TBL	Total bilirubin
TEAE	Treatment emergent adverse event
ULN	Upper limit of the normal range
WHODRUG	World Health Organization Drug Dictionary

2. INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to describe the planned analyses and data displays to be included in the Clinical Study Report (CSR) for Protocol AT-03A-017.

Protocol Revision Chronology:		
Protocol	08-SEP-2022	Original (Version 1.0)
Protocol Amendment	15-JUN-2023	Amendment (Version 2.0)

The SAP is based on:

- Protocol No. AT-03A-017 Amendment, Version 2.0, dated June 15, 2023
- IND 149508 FDA End of Phase 2 Meeting Minutes, issued 18 AUG 2022
- ICH E9: Statistical Principles for Clinical Trials[1].

The purpose of this document is to provide details on study populations and on how the variables will be derived, how missing data will be handled as well as details on statistical methods to be used to analyze the safety and efficacy data in Study AT-03A-017.

The document may evolve over time, for example, to reflect the requirements of protocol amendments or regulatory requests. However, the final SAP will be finalized, approved by the Sponsor, and placed on file before unblinding of study data. Deviations from the final approved plan will be noted in the clinical study report.

Study AT-03A-017 will have evaluations of efficacy and safety in 2 strata: a supportive-care only group that will receive either bemnifosbuvir (BEM) or placebo without any additional COVID-19 antiviral drugs and a combination antiviral group that will receive either BEM or placebo in addition to an antiviral COVID-19 therapy. The primary analysis is based on subjects in the supportive-care only stratum and this stratum will comprise the majority of subjects enrolled into the study. [redacted]
Unless otherwise specified, all analyses will be carried out on each stratum. Further information can be found in the protocol.

SAP version 5 includes additional subgroup analysis, and clarifications from previous SAP versions for programming.

3. STUDY OBJECTIVES AND ENDPOINTS

3.1. Study Objectives

3.1.1. Primary Objective

The primary objective is to evaluate the efficacy of bemnifosbuvir (BEM) compared with placebo in reducing all-cause hospitalization or all-cause death in COVID-19 outpatients receiving only supportive care.

3.1.2. Secondary Objectives

The following are secondary objectives:

- To evaluate the efficacy of BEM compared with placebo on:
    - All-cause hospitalization, all-cause death, or COVID-19 related medically attended visits through Day 29 and Day 60
    - COVID-19 related hospitalization or all-cause death through Day 29 and Day 60
    - All-cause death through Day 29 and Day 60
    - COVID-19 related complications (eg, death, hospitalization, radiologically confirmed pneumonia, acute respiratory failure, sepsis, coagulopathy, pericarditis/myocarditis, cardiac failure) through Day 29 and Day 60
    - COVID-19-related medically attended visits (hospitalization, emergency room (ER) visit, urgent care visit, physician's office visit, or telemedicine visit) or all-cause death through Day 29 and Day 60
    - COVID-19 related medically attended visits through Day 29 and Day 60
    - COVID-19 symptom relapse
- To evaluate the antiviral activity of BEM compared with placebo on viral load rebound
- To evaluate the safety of BEM compared with placebo.

3.1.3. Exploratory Objectives

The exploratory objectives are

- To evaluate the efficacy of BEM compared with placebo on:
    - Sustained alleviation or improvement of COVID-19 symptoms
    - Return to usual (pre-COVID-19) health
    - Requirement for post-randomization rescue therapy
    - The antiviral activity of BEM compared with placebo
- To assess plasma concentrations of AT-511 (free base form of AT-527 [BEM]) and its metabolites from sparse PK sampling at steady-state.

3.2. Study Endpoints

Estimands corresponding to primary and secondary endpoints are defined in Table 2 .

3.2.1. Primary Endpoints

The primary efficacy endpoint is the proportion of subjects in the supportive-care-only population who are hospitalized for any cause or died due to any cause through Day 29.

3.2.2. Secondary Endpoints

The secondary endpoints are:

Efficacy:

- Proportion of subjects with all-cause hospitalization or all-cause death through Day 60
- Proportion of subjects with all-cause hospitalization, all-cause death, or COVID-19 related medically attended visits through Day 29 and Day 60
- Proportion of subjects with COVID-19-related hospitalization or who died due to any cause death through Day 29 and Day 60
- Proportion of subjects who died due to any cause through Day 29 and Day 60
- Proportion of subjects with COVID-19-related complications (eg., death, hospitalization, radiologically confirmed pneumonia, acute respiratory failure, sepsis, coagulopathy, pericarditis/myocarditis, cardiac failure) through Day 29 and Day 60
- Proportion of subjects with COVID-19-related medically attended visits (hospitalization, emergency room (ER) visit, urgent care or physician's office visit or telemedicine visit) or who died due to any cause through Day 29 and Day 60
- Proportion of subjects with COVID-19 symptom relapse through Day 29
- Proportion of subjects with viral load rebound through Day 29.

Safety:

- The incidence and severity of adverse events (AEs) and serious adverse events (SAEs).

The primary endpoint will also be assessed in the combination antiviral population as a secondary endpoint. All secondary endpoints will be assessed in both the supportive-care-only population and the combination antiviral population.

3.2.3. Exploratory Endpoints

Efficacy:

- Time to sustained alleviation or improvement of COVID-19 symptoms
- Time to self-reported return to usual (pre-COVID-19) health
- Proportion of subjects requiring post randomization rescue therapy

Note: For subjects randomized into the combination antiviral population, antiviral medications administered post-baseline within 24 hours of first study drug, as part of the combination regimen, are not considered rescue medication.

Note: Rescue therapy for the purpose of the endpoint includes only treatments with direct activity against SARS-CoV-2; and that is administered for COVID-19-related clinical decline.

- Duration and severity of COVID-19 symptom relapse
- Proportion of subjects with both COVID-19 symptom relapse and viral rebound
- Change from baseline in amount of SARS-CoV-2 viral RNA as measured by RT-qPCR overtime
- Time to cessation of SARS-CoV-2 viral shedding by RT-qPCR
- Proportion of subjects negative for SARS-CoV-2 viral RNA by RT-qPCR over time
- Area under the curve (AUC) in the amount of SARS-CoV-2 viral RNA as measured by RT-qPCR
- Potential future biomarker analyses[redacted]
- Pharmacokinetics: Plasma concentration of AT-511 (free base form of AT-527 [BEM]) and its metabolites

[redacted]

Table 2: Estimand Definitions for Primary and Key Secondary Endpoints

Estimand	Endpoint	Intercurrent event handling	Population-Level Summary for the Variable	Handling of Missing Data: Study Withdrawal/ Lost to Follow-Up
Primary: difference in the proportion of subjects in the supportive-care population with all-cause hospitalization or death by Day 29 and Day 60, regardless of intercurrent events	The proportion of subjects in the supportive-care-only population who are hospitalized for any cause or died due to any cause through Day 29	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy.	The common risk difference	The outcome measure up to point of study withdrawal will be used
Secondary: difference in the proportion of subjects in the supportive-care population with all-cause hospitalization, all-cause death, or COVID-19 related medically attended visits through Day 29 and Day 60, regardless of intercurrent events	The proportion of subjects in the supportive-care population who are hospitalized for any cause or died due to any cause, or had COVID-10 related medically attended visits through Day 29 and Day 60	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy.	The common risk difference	The outcome measure up to point of study withdrawal will be used

Estimand	Endpoint	Intercurrent event handling	Population-Level Summary for the Variable	Handling of Missing Data: Study Withdrawal/ Lost to Follow-Up
Secondary: difference in the proportion of subjects in the combination antiviral population with all-cause hospitalization or death by Day 29 and Day 60, regardless of intercurrent events	The proportion of subjects in the combination antiviral population who are hospitalized for any cause or died due to any cause through Day 29	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy.	The common risk difference	The outcome measure up to point of study withdrawal will be used
Secondary: difference in the proportion of subjects in the supportive-care-only population with COVID-19 related hospitalization or any cause death by Day 29 and Day 60, regardless of intercurrent events	The proportion of subjects in the supportive-care-only population with COVID-19-related hospitalization or who died due to any cause death through Day 29 and Day 60	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy.	The common risk difference	The outcome measure up to point of study withdrawal will be used

Estimand	Endpoint	Intercurrent event handling	Population-Level Summary for the Variable	Handling of Missing Data: Study Withdrawal/ Lost to Follow-Up
Secondary: difference in the proportion of subjects in the supportive-care-only population with COVID-19-related complications by Day 29 and Day 60, regardless of intercurrent events	The proportion of subjects in the supportive-care-only population with COVID-19-related complications through Day 29 and Day 60	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy.	The common risk difference	The outcome measure up to point of study withdrawal will be used
Secondary: difference in the proportion of subjects in the supportive-care-only population with COVID-19-related medically attended visits or any cause death by Day 29 and Day 60, regardless of intercurrent events	The proportion of subjects in the supportive-care-only population with COVID-19-medically attended visits or who died due to any cause through Day 29	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy.	The common risk difference	The outcome measure up to point of study withdrawal will be used

Estimand	Endpoint	Intercurrent event handling	Population-Level Summary for the Variable	Handling of Missing Data: Study Withdrawal/ Lost to Follow-Up
Secondary: difference in the proportion of subjects in the supportive-care-only population who died, regardless of intercurrent events through Day 29 and Day 60	The proportion of subjects in the supportive-care-only population who died due to any cause through Day 29 and Day 60	Treatment-policy strategy approach will be taken. Patients are followed, assessed and analyzed irrespective of their completion of the planned course of the study treatment and irrespective of rescue and concomitant therapy. Death: composite strategy, considered as event	The common risk difference	The outcome measure up to point of study withdrawal will be used
Secondary: difference in the proportion of subjects in the supportive-care-only population with COVID-19 symptom relapse by Day 29, regardless of intercurrent events	Proportion of subjects with COVID-19 symptom relapse in the supportive-care-only population through Day 29	Treatment discontinuation or rescue treatment prior to event: treatment-policy strategy approach will be taken.	The common risk difference	The outcome measure up to point of study withdrawal will be used

Estimand	Endpoint	Intercurrent event handling	Population-Level Summary for the Variable	Handling of Missing Data: Study Withdrawal/ Lost to Follow-Up
Secondary: difference in the proportion of subjects in the supportive-care-only population with viral load rebound by Day 29, regardless of rescue medication	The proportion of subjects in the supportive-care-only population with viral load rebound through Day 29	Treatment discontinuation or rescue treatment prior to event: treatment-policy strategy approach will be taken.	The common risk difference	The outcome measure up to point of study withdrawal will be used

4. STUDY DESIGN

4.1. Summary of Study Design

AT-03A-017 is a Phase 3 randomized, double-blind, placebo-controlled study, conducted in non-hospitalized adult subjects (≥18 years of age) with mild or moderate COVID-19.

Age and comorbidities will determine a patient's risk and eligibility. Onset of COVID-19 symptoms will have occurred ≤5 days before randomization.

BEM 550 mg or matching placebo will be administered BID for 5 days (10 doses, with each dose including 2 tablets).

Locally available standard of care (SOC) treatment is to be initiated in parallel with study drug (add-on). SOC treatments with direct antiviral activity against SARS-CoV-2 are allowed, but are to be initiated during screening and/or within 24 hours prior to or after randomization.

Subjects will be randomized 1:1 (active:placebo) to one of the following arms:

- 550 mg BEM (2 x 275 mg tablets) BID for 5 days
- Placebo (2 tablets) BID for 5 days

Randomization will be stratified by geographic region (North America, Europe, Asia, and rest of world [ROW]), age (< 65 years old vs ≥ 65 years old) and by the type of SOC which the investigator plans to administer in parallel at time of study drug initiation. The type of SOC received will be used to classify subjects into two categories:

- Combination antiviral stratum (subsequently referred to as the combination antiviral population): In addition to study drug, the subject will be initiated on a local SOC regimen which includes a treatment with direct SARS-CoV-2 antiviral activity. Examples include nirmatrelvir/ritonavir, molnupiravir, favipiravir, mAbs
- Supportive-care-only stratum (subsequently referred to as the supportive-care-only population): the SOC regimen the subject will receive does not include any treatment with direct SARS-CoV-2 antiviral activity

Although randomization is stratified by whether COVID-19 drugs with antiviral effects are administered, both strata allow local SOC therapies without direct antiviral effects (i.e., corticosteroids).

Patients who require SOC treatment with a non-compatible antiviral treatment, defined as those with a potential drug-drug interaction (DDI) with BEM (e.g., remdesivir), are excluded.

For subjects who deteriorate on study, any rescue therapy which is SOC for progression to severe COVID-19 is permitted. If rescue therapy will be remdesivir, study drug will be discontinued before initiation of remdesivir.

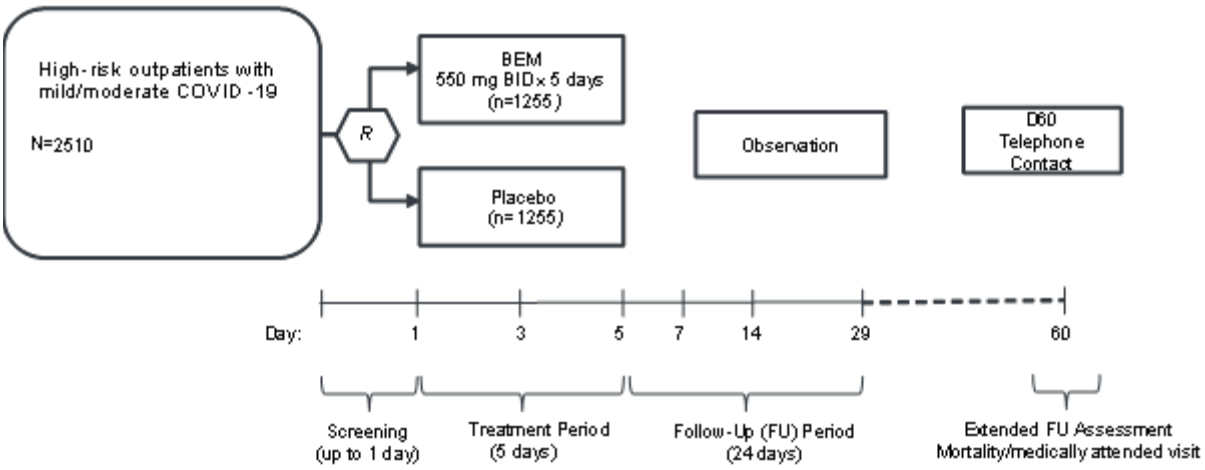
Overall, up to 2510 eligible subjects with COVID-19 will be enrolled, with a target of 2210 subjects in the supportive-care-only population and 300 subjects in the combination antiviral population. The primary study endpoint (hospitalization/death) will be evaluated in the supportive-care-only population.

The primary study period is up to 30 days, including up to 1 day for screening and 29 days following randomization. Extended follow-up data will be collected through Day 60 for

mortality and medically attended visit (MAV) assessment. The Day 60 follow-up contact marks the end of study participation for each subject and the study ends when the last subject completes the Day 60 follow-up contact.

A schematic of the study design is presented in Figure 1.

Figure 1: Study Design



4.2. Definition of Study Drugs

Study drug (550 mg BEM or matching placebo) is to be administered twice daily for 5 days, i.e., a total of 10 doses (20 tablets total). Each dose will be taken as 2 tablets (2x 275 mg BEM tablets or 2 matching placebo tablets).

4.3. Sample Size Considerations

The primary endpoint of proportion of subjects with all-cause hospitalization and all-cause mortality through Day 29 will be compared between treatments using a Cochran-Mantel-Haenszel (CMH) test adjusted for the stratification factors of age and region. The overall type I error will be controlled at a one-sided 0.025 level. Total beta spending was specified at a 0.10 level.

The study was planned to have interim analyses at 30% (futility only) and 65% (futility, safety and efficacy) of the originally planned number of subjects. The interim analyses were removed due to fast enrollment such that the final sample size was reached prior to the first planned interim analysis.

4.3.1. Sample Size Justifications

Assuming the combined all-cause hospitalization and all-cause mortality rate will be 4% in the placebo arm, the sample size needed to detect a true 60% reduction (corresponding to an absolute difference in rates of 2.4%) in the BEM treatment arm with 90% power at a 1-sided significance

level of 2.5% was determined to be 1982 randomized participants in the supportive-care-only population. Sample size was calculated using SAS (version 9.4) for a 2-sample proportion Chi-square test. Considering the planned interim analyses with both alpha and beta spending according to Lan-Demets approximation of O'Brien Fleming method the sample size should be inflated by a multiplier of 1.055 to a target sample size of 2100 participants (see Section 11.1). Assuming 95% of randomized subjects who receive only supportive care as SOC will meet the mITT-S population criteria (see Section 6.5.4), the study targets total enrollment of 2210 subjects in the supportive-care-only population. If the true rate in the placebo arm is 3%, and assuming a reduction of 60% with BEM, the planned study has at least 80% power to detect a treatment effect. Planned enrollment in the combination arm is up to 300 subjects.

4.4. Randomization

Subjects will be randomized 1:1 (active:placebo) to one of the following arms:

- 550 mg BEM (2 x 275 mg tablets) BID for 5 days
- Placebo (2 tablets) BID for 5 days

Randomization will be stratified by geographic region (North America, Europe, Asia, and ROW), age (< 65 years old, ≥ 65 years old) and by the type of SOC which the investigator plans to administer in parallel at time of study drug initiation. The primary stratification factor is the type of background SOC that the investigator plans to administer as this determines whether the subject is in the supportive-care only group (i.e. may be included in the primary analysis population) or the combination antiviral group. Within each of the primary strata, randomization is further stratified by geographic region (North America, Europe, Asia, and ROW) and age (<65 and ≥65).

The type of SOC received will be used to classify subjects into two categories:

- Combination antiviral stratum (subsequently referred to as the combination antiviral population):In addition to study drug, the subject will be initiated on a local SOC regimen which includes a treatment with direct SARS-CoV-2 antiviral activity. Examples include nirmatrelvir/ritonavir, molnupiravir, favipiravir, mAbs
- Supportive-care-only stratum (subsequently referred to as the supportive-care-only population): the SOC regimen the subject will receive does not include any treatment with direct SARS-CoV-2 antiviral activity

4.5. Clinical Assessments

The following assessments will be performed at Screening, defined as within 24 hours of first dose of study drug, only:

- Demographics
- Medical history and baseline characteristics
- Pre-existing symptoms
- Height and weight
- Complete physical examination

- Pregnancy test (females of childbearing potential)
- Rapid test for influenza.

Information on study drug dosing will be collected on Days 1-6 and a plasma sample for evaluation of sparse pharmacokinetics will be collected on Day 5.

The following assessments will be performed at visits corresponding to Days 1, 5, 14 and Early Termination (ET):

- Central safety laboratory values
- Potential blood biomarkers [redacted]

The following assessments will be collected at each study visit from Day 1 onward (during the 29-day follow-up period i.e., Days 1, 3, 5, 7, 14, 29, and ET if applicable) with exceptions noted:

- Vital signs, to Day 14
- Symptom-targeted physical examination, as clinically indicated
- Nasopharyngeal (NP) swab for central assessment of SARS-CoV-2 virus (quantitative RT-PCR and potential viral titer)
- Adverse events (SAEs out to Day 60)
- Concomitant medications
- Assessment of COVID-19 related medically attended visit(s) (out to Day 60).

Patient reported COVID-19 symptoms will be collected daily on Days 1-29. Symptoms are captured via the COVID-19 symptom diary (see Section 11.1).

At the extended follow-up visit on Day 60, planned collection includes assessment of medically attended visit(s) and serious adverse events.

5. PLANNED ANALYSES

5.1. Interim Analyses

A Data and Safety Monitoring Board (DSMB) will be responsible for ongoing monitoring of safety for the study. The DSMB members are independent of the study team and are external to the Sponsor. Unblinded analyses will be prepared and provided to the DSMB by an independent statistical group at the contract research organization (CRO). A DSMB charter describes the role of the DSMB in more detail.

Efficacy and futility interim analyses were planned at the beginning of this study but were removed due to fast enrollment such that the final sample size was reached prior to the first planned interim analysis.

5.2. Final Analyses

6. FINAL ANALYSES WILL BE PERFORMED AFTER DATABASE LOCK WHEN ALL SUBJECTS HAVE COMPLETED OR DISCONTINUED FROM THE STUDY AND DATA CLEANING HAS BEEN COMPLETED. GENERAL CONSIDERATIONS FOR DATA ANALYSES AND HANDLING

6.1. General Summary Table and Individual Subject Data Listing Considerations

Unless otherwise specified, continuous variables will be summarized by using the number of non-missing observations, arithmetic mean, standard deviation (SD), minimum, first quartile, median, third quartile and maximum values as summary statistics; categorical variables will be summarized by using the frequency count and the percentage of subjects in each category as summary statistics. Descriptive summaries will be tabulated by treatment group.

When a summary tabulates the number of subjects, subjects will only be counted once for a category (e.g., a particular adverse event type, a particular adverse event term, a particular concomitant medication) unless otherwise specified.

All statistical tests will be performed at a 0.025 significance level using one-sided tests, unless otherwise noted. Statistical tests will use stratification factors (age and region) according to randomization. Stratification factors may need to be removed or have their levels combined so that low or zero counts in one stratum do not invalidate relevant statistical tests.

Unless otherwise specified, all efficacy and virology analyses will be performed on the modified intent-to-treat (mITT) analysis populations for each stratum. All safety analyses, demographics, and baseline characteristics will be performed on the safety population. Refer to Section 6.5 for details on analysis population definitions.

All Safety, Efficacy, and Tolerability data will be listed. In cases where additional information may be helpful, considerations for listings will be noted in this document.

6.2. General Post Text Summary Table and Individual Subject Data Listing Format Considerations

By-subject listings will be sorted by population stratum (i.e. supportive-care only, combination antiviral), antivirals used (for combination antiviral only), treatment, subject ID number, assessment/start date, and assessment/start time (where applicable). Subject identification will include age, sex, country (as 3-digit code) and race.

6.3. Data Management

Data management and cleaning will be conducted in accordance with the standard operating procedures (SOP) of the CRO for the study with oversight in accordance with the Sponsor's corresponding SOP. Derived datasets, data analyses and summary tables will be created using SAS® software, version 9.4 or above.

6.4. Data Presentation Conventions

Unless otherwise indicated, data summaries and individual listings will be presented by population stratum and treatment. Treatment will be considered as BEM and placebo within each population stratum. For presentations of efficacy analysis, subjects will be grouped according to the treatment to which they were randomized. For presentations on the safety analysis, subjects will be grouped according to the treatment that was received.

6.5. Analysis Populations

6.5.1. Screen Failures

Screen failures will include subjects who fail to meet inclusion/exclusion (I/E) criteria. Subjects who are enrolled and do not fail any I/E criteria but decide not to participate in the study prior to randomization will also be included here.

6.5.2. Safety Population

The safety population will include all subjects who received at least one dose of the study drug (BEM or placebo). The study population will be analyzed as treated. Any subjects who were randomized more than once will be excluded from the summary tables, but the data will be included in listings and indicated appropriately.

6.5.3. ITT Population

The intent-to-treat (ITT) population will include all randomized subjects and will be analyzed as randomized.

6.5.4. mITT Population

The modified intent-to-treat (mITT) population includes all randomized subjects who received ≥ 1 dose of study drug and have ≥ 1 positive SARS-CoV-2 result (from NP swab samples via central testing) through Day 5. The majority of subjects are expected to have the positive SARS-CoV-2 qualifying sample at baseline, but the window is extended to allow the inclusion of subjects who may have a baseline false negative (via central testing) or missing sample at baseline. Any subjects who were randomized more than once will be excluded from the summary tables, but the data will be included in listings and indicated appropriately.

The modified intent-to-treat-S (mITT-S) population includes all randomized subjects who received ≥ 1 dose of study drug, have ≥ 1 positive SARS-CoV-2 result (from NP swab samples via central testing) through Day 5, and who are stratified into the supportive-care-only population according to the standard of care received (i.e., per EDC).

The modified intent-to-treat-C (mITT-C) population includes all randomized subjects who received ≥ 1 dose of study drug, have ≥ 1 positive SARS-CoV-2 result (from NP swab samples via central testing) through Day 5, and who are stratified into the combination antiviral population according to the standard of care received (i.e., per EDC).

The mITT populations are further restricted from the ITT population, but the intent-to-treat principle has been maintained as the requirements only define the target population while not introducing bias between the two treatment arms. The mITT-S and mITT-C populations are

displayed separately to present the treatment comparisons in two different standard of care scenarios. If a subject is mis-stratified in randomization, actual stratification will be used in the analysis performed by strata.

6.5.5. Per Protocol Population

The per-protocol population for the supportive care population is a subset of the mITT-S population that:

- Received all 10 doses of study drug as per protocol (with the exception of those subjects who discontinued study drug to initiate remdesivir rescue therapy, or discontinued study due to hospitalization or death)
- Did not have any inclusion/exclusion criteria protocol violations.

Subjects are excluded from the mITT-S per-protocol population if they received non-study drugs with SARS-CoV-2 antiviral activity (e.g. nirmatrelvir/ritonavir, molnupiravir, favipiravir, monoclonal antibodies (mAbs)) for the current SARS-CoV-2 infection from screening (or 24 hours prior to randomization) through Day 29 (with the exception of non-study drugs administered as rescue antiviral therapy).

6.5.6. Completers

Subjects considered as having completed the primary study follow-up are those in the mITT population who were followed through to at least Day 29 or had a hospitalization or death event prior to Day 29. Any subject who has a Day 29 (or subsequent scheduled) visit or an assessment on a date corresponding to study day 29 or later (see Section 6.7.2) will be considered as having been followed through to at least Day 29.

6.5.7. PK population

The PK population will include all subjects who receive at least 1 dose of AT-527 and have at least 1 pharmacokinetic measurement.

6.6. Baseline Definition

Baseline is defined as the last available assessment prior to first dose of study treatment (BEM or placebo). For analyses of virology data, if there is no pre-study drug assessment available then baseline can be up to 4 hours post first dose, with the earliest assessment within this time window used as baseline in the case of multiple qualifying assessments. For analysis of symptoms, Day 1 assessment will be used as the baseline.

6.7. Derived and Transformed Data

6.7.1. Baseline Age

Baseline age is the age at the time of randomization.

6.7.2. Study Day

Study day will be calculated from the first dosing date of study drug (FDD) and derived as follows:

- Assessment date ≥ date of first study drug: Assessment Date – FDD + 1
- Assessment date < date of first study drug: Assessment Date – FDD.

If a subject is randomized but does not receive any study drug, the randomization date will be used for reference, in place of FDD.

Relative symptom start day will be calculated from the symptom start date and derived as follows:

- FDD – symptom start date +1

6.7.3. Change from Baseline

Where applicable, change from baseline will be computed simply as the difference between the post-baseline value and the baseline value.

6.7.4. Visit Windows

For the purposes of by-visit analyses for scheduled clinical assessments (i.e., excluding patient-reported assessments), visit windows for inclusion in the analysis are defined in Table 3 . Note that the definition of baseline is provided in Section 6.6.

Table 3: Visit Windows (Days)

Visit	Target Study Day	Visit Window, Study Day(s)
Baseline/Day 1[a]	1	1
Day 3	3	2 – 4
Day 5	5	4 – 6
Day 7	7	7 – 10
Day 14	14	11 – 21
Day 29	29	22 - 36
Day 60	60	>36
a. If a Day 1 assessment is not available, a qualifying baseline assessment can be utilized.		

Patient reported outcomes are to be reported daily on Days 1 – 29 and data will be based on the corresponding day. If there is more than one assessment on a day, then the least favorable assessment will be utilized for analysis.

There is overlap at study day 4 for the Day 3 and Day 5 windows to increase the chance to have qualifying observations at both of these timepoints. In the case where a measurement is captured on Day 4, the assignment for the observation will be according to the following:

- If there is another measurement on study day 5 and:
    - Another qualifying measurement for Day 3: day 4 measurement not assigned to Day 3 or Day 5
    - Otherwise, day 4 measurement assigned to day 3 and day 5 measurement assigned to Day 5.
- Else if there is another measurement on study day 6 and:
    - Another qualifying measurement for Day 3: day 4 measurement not assigned to Day 3 or Day 5
    - Otherwise, day 4 measurement assigned to Day 3 and day 6 measurement assigned to Day 5.
- Else day 4 measurement is assigned to Day 5.

6.7.5. Multiple Assessments

If multiple measurements are obtained within a visit window for a subject, the value obtained closest to the target timepoint will be used. For any visit, no more than one value per subject will be included in an analysis. In the event of multiple observations that are equidistant from, but not on, the target day (i.e., observations made in a similar time frame before and after target), the later value, based on date and time, will be selected.

6.8. Handling of Missing Data

6.8.1. Missing Efficacy Endpoints

Missing efficacy data will not be imputed unless otherwise specified in the specific section describing an endpoint. Sensitivity analyses will be carried out to evaluate the impact of missing efficacy data.

6.8.2. Missing Start and Stop Dates for Prior and Concomitant Medication

Where the start date for a concomitant medication is missing, the medication will be assumed to have started prior to enrolment in the study. Where the stop date for a concomitant medication is missing, the medication will be assumed to be ongoing to the date of the subject's last assessment.

6.8.3. Missing Start and Stop Dates for Adverse Events

Where the entire start date for an adverse event is missing, the adverse event will be assumed to be treatment-emergent. In this case, if the duration of an adverse event is computed, the start date will be considered as the date associated with Day 1. If the month and year of an adverse event is provided, then the start date will be considered as the later date between the first day of the month/year and the treatment start date.

Where the stop date for an adverse event is missing, the adverse event will be assumed to be ongoing to the date of the subject's last assessment. If the month and year of the stop date is provided, then the stop date will be considered as the earlier date between the last day of the month/year and the last assessment date.

6.8.4. Missing Numeric Viral Load Assessment

Where SARS-CoV-2 RNA is detected but the numeric value is "<500", it will be assigned as LLoQ-1 (i.e., 499). If SARS-CoV-2 RNA is not detected, it will be imputed as LOD – 1 (i.e., 249). If SARS-CoV-2 RNA is ">500000000", it will be assigned as 500000001 for the purpose of quantitative summaries. For samples reported outside the limits of quantification (e.g.., '<500', or '>500000000'), listings will display "<500" or ">500000000", as appropriate.

6.8.5. Missing COVID-19 symptom diary assessment

If a subject was missing all symptom diary assessments, this subject will be excluded from analysis. For the time-to-event analyses, missing individual symptom assessments at times will be imputed prior to statistical analysis. If there is missing data between two assessments, the missing data will be imputed as the worst outcome of the two. Missing data after the last recorded assessment will not be imputed; instead censoring may be utilized for the time to event endpoint in cases where the endpoint was not achieved. For time to sustained alleviation/improvement of symptoms, censoring will be applied thus:

- if the last observed time point does not meet alleviation/improvement criteria, the result is censored at the time of the last recorded result
- if the last observed time point meets alleviation/improvement criteria but the requirement for sustainment (i.e. criteria not maintained over 4 day period) has not yet been met, the result is censored using the approach described in Section 8.8.

Missing baseline assessments will be imputed according to the following rules: 1) missing individual symptoms and worst severity of COVID-19 symptoms will be imputed as 'mild'. 2) Missing assessments for "loss of taste" and "loss of smell" will be imputed as 'no'. 3) Missing general physical health will be imputed as "fair". 4) Missing returned to usual health will be imputed as "no". 5) Missing symptom relapse will be imputed as "not applicable".

7. STUDY POPULATION

7.1. Subjects Disposition

The following will be tabulated (number and percentage) by treatment group for the ITT population:

- Received study drug as randomized
- Did not receive study drug as randomized, overall and by reason (e.g., not treated, incorrect treatment received)
- Completed treatment
- Discontinued treatment, overall and by reason
- Completed assessment through Day 29
- Did not complete assessment through Day 29, overall and by reason
- Completed extended follow-up (i.e., to Day 60).
- Did not complete extended follow-up, overall and by reason.

The following will also be tabulated (number and percentage) by treatment for the ITT population:

- Subjects randomized and treated in each country and investigator site
- Subjects randomized and treated by each randomization stratum
- Subjects included and excluded (with reason) for each analysis population.

7.2. Screen Failures

The numbers and proportions of subjects who enrolled but were not randomized will be presented as:

- Screen failures who were not randomized
- Subjects who do not have indication of failing to meet I/E criteria but were not randomized.

Screen failure reasons will also be tabulated. A subject may have more than one reason for screen failure and all reasons for a subject will be counted, i.e., the sum of subjects across screen failure reasons may exceed the total number of screen failures.

7.3. Protocol Deviations

A final identification and classification of important protocol deviations will be made prior to the final unblinding of the database and will be documented, in accordance with the study's Protocol Deviations Management Plan. The number and percentage of subjects with important protocol deviations and the total number of important protocol deviations by deviation reason (e.g., nonadherence to study drug, violation of select inclusion/exclusion criteria, lab assessment, randomization, or study procedures, etc.) will be summarized by treatment for the ITT

population. A by-subject listing will be provided for those participants with important protocol deviations.

7.4. Demographic and Baseline Characteristics

To review treatment group comparability within the study, a number of variables collected at baseline will be presented by treatment groups on the mITT-S, mITT-C and safety analysis populations. No statistical comparisons will be made on these characteristics.

Demographic variables to be summarized include:

- Age
- Age group (18-49, 50-64, 65 - 79 , ≥ 80)
- Sex
- Race
- Ethnicity
- Country
- Region
- Weight
- Body mass index (BMI).

Baseline Disease Characteristics to be summarized include:

- SARS-CoV-2 RT-PCR viral load at baseline
- SARS-CoV-2 IgG seropositive status at baseline
- Presence of individual high-risk factors (age groups, obesity by age groups, diabetes by age groups, cardiovascular disease by age groups, chronic lung disease by age groups, immunocompromised or immunosuppressed by age groups, other high-risk conditions by age groups)
- Time from symptom onset to study treatment in days
- COVID-19 vaccination status
- Baseline lineage

7.5. Medical History and Medical Conditions Present at Entry

Complete medical history will be obtained at screening. Medical history records will be coded, using the current version of Medical Dictionary for Regulatory Activities (MedDRA), and summarized descriptively using the safety population.

7.6. Prior Medication History and Medications Present at Entry

Medications will be coded using the most recent version of the World Health Organization Drug Dictionary (WHODRUG). Prior medications will be summarized descriptively by treatment group for the safety population. Previous treatments that have been stopped prior to date of first study drug dose will be summarized separately from medications that were started and continued into the study drug treatment period.

A summary of prior medications with indication of COVID-19 will also be provided.

8. EFFICACY

8.1. General Considerations

Where adjustment for stratification factors (age and region) is specified, the group will be based on the randomization assignment. All efficacy endpoints will be evaluated post treatment.

8.2. Testing Statistical Assumptions Including Comparability at Baseline

No testing will be carried out on baseline characteristics.

The primary analysis method for comparison of treatment groups is the Cochran-Mantel-Haenszel (CMH) test. For CMH, small cell counts could be problematic. The Mantel-Fleiss criterion will be used to assess the validity of the chi-square approximation for the distribution of the Mantel-Haenszel statistic in the analyses. In the event there is indication that the test statistic may not be valid, strata will be removed from the analysis until the criterion for validity is met. Any removal of strata will be carried out in the following order: age first, then region. If all strata are indicated to be removed, then an unstratified comparison will be carried out.

8.3. Statement of the Null and Alternate Hypotheses

The null hypothesis is that the likelihood of all-cause hospitalization or death to Day 29 is the same between treatment groups in the supportive care population. The likelihood of the event is estimated by the proportion of subjects with the event, i.e., the primary endpoint.

The alternative hypothesis is that the treatment groups differ in the likelihood of all-cause hospitalization or death to Day 29, with the rate being lower in the BEM arm as the one-sided (0.025 level) alternative of interest.

8.4. Subgroup Analyses

To examine the robustness of the treatment effect, estimates of the difference between treatments in the proportions of subjects in the supportive-care-only population who are hospitalized for any cause or who died due to any cause through Day 29 (i.e., the primary endpoint) will be computed with corresponding 95% confidence intervals, per Wilson (Score) method without continuity correction, for the subgroups listed below. Estimates with confidence intervals will be presented in forest plots.

- Geographic region (North America, Europe, Asia, ROW)
- Age (<65, ≥ 65 years)
- Sex (female, male)
- Race
- Time from onset of symptoms (≤3 days, >3 days)
- Baseline serology status (positive, negative)
- Baseline viral lineage (as applicable)
- COVID-19 vaccination status

- Age (<80, ≥80)
- BMI (<30, ≥ 30)
- Diabetes mellitus (yes, no)
- Cardiovascular disease or hypertension (yes, no)
- Chronic lung disease (yes, no)
- Immunocompromised/immunosuppressed (yes, no)
- Baseline viral load (<median, >=median).

8.5. Multiple Comparisons and Multiplicity

If a statistically significant result is found on the test of the primary endpoint, sequential testing will be performed for the secondary endpoints in the following order. When endpoint is tested for both Day 29 and Day 60, the order will be Day 29 first, followed by D60.

1. Proportion of subjects in the supportive-care-only population with all-cause hospitalization or all-cause death through Day 60
2. Proportion of subjects in the supportive-care-only population with all-cause hospitalization, all-cause death, or COVID-19 related medically attended visits through Day 29 and Day 60.
3. Proportion of subjects in the supportive-care-only population with COVID-19-related hospitalization or who died due to any cause death through Day 29 and Day 60
4. Proportion of subjects in the supportive-care-only population who died due to any cause through Day 29 and Day 60
5. Proportion of subjects in the supportive-care-only population with COVID-19-related complications (e.g., death, hospitalization, radiologically confirmed pneumonia, acute respiratory failure, sepsis, coagulopathy, pericarditis/myocarditis, cardiac failure) through Day 29 and Day 60
6. Proportion of subjects in the supportive-care-only population with COVID-19-related medically attended visits (hospitalization, ER visit, urgent care or physician's office visit or telemedicine visit) or who died due to any cause through Day 29 and Day 60.
7. Proportion of subjects in the supportive-care-only population with COVID-19-related medically attended visits through Day 29 and Day 60
8. Proportion of subjects in the supportive-care-only population with COVID-19 symptom relapse through Day 29
9. Proportion of subjects in the supportive-care-only population with viral load rebound through Day 29
10. Proportion of subjects in the combination-antiviral population with COVID-19 symptom relapse through Day 29
11. Proportion of subjects in the combination-antiviral population with viral load rebound through Day 29

8.6. Analysis of the Primary Efficacy Endpoint

Table 4: Summary of Analyses Related to the Primary Efficacy Endpoint

Analysis	Population	Handling of Intercurrent Events	Analysis Approach
Primary Analysis: All-cause hospitalizations and deaths (through Day 29)	mITT-S	Treatment policy	Cochran-Mantel-Haenszel
Supplemental Analysis: All-cause hospitalizations and deaths (through Day 29)	mITT-S Completers	Treatment policy	Cochran-Mantel-Haenszel
Supplemental Analysis: All-cause hospitalizations and deaths (through Day 29)	mITT-S Per-protocol	Treatment policy	Cochran-Mantel-Haenszel
Supplemental Analysis: All cause hospitalizations and deaths (through Day 29)	mITT-S	Intervention with COVID-19 rescue as event (composite handling); Treatment policy otherwise	Cochran-Mantel-Haenszel
Supplemental Analysis: Time to all cause hospitalization and death (through Day 29)	mITT-S	Intervention with COVID-19 rescue as event (composite handling); Treatment policy with censoring for early study discontinuation (prior to Day 29) without event	Kaplan-Meier plot, stratified logrank test

8.6.1. Primary Efficacy Analysis

The primary efficacy objective of this study is to evaluate the efficacy of BEM based on the proportion of subjects who are hospitalized for any cause or who died due to any cause through Day 29 in the supportive-care-only population. Hospitalization is defined as >24 hours of acute care, in a hospital or similar acute care facility, including Emergency Rooms or temporary facilities instituted to address medical needs of those with severe COVID-19 during the COVID-19 pandemic. This includes specialized acute medical care unit within an assisted living facility

or nursing home. This does not include hospitalization for the purposes of public health and/or clinical trial execution. Furthermore, scheduled hospitalization for a pre-planned procedure (including chemotherapy) to treat a pre-existing condition that has not worsened is not an SAE, and would not be considered a hospitalization for the purposes of the efficacy analyses. A pre-existing condition is a clinical condition that is diagnosed prior to the use of study drug and is documented in the subject's medical history.

The primary estimand is the difference in the proportions of subjects who are hospitalized for any cause or who died due to any cause through Day 29 in the supportive-care-only population, and the primary treatment effect comparison will be between the BEM arm and the placebo arm. This proportion will be compared between the BEM and placebo arms using a CMH test adjusted for the stratification factors of age and region. The CMH p-value and estimated common risk difference, odds ratio and relative risk, along with the corresponding 95% confidence intervals (CIs) for these estimates, will be presented. The confidence interval for the common risk difference will be computed using the Klingenberg method[5]. Summary tables will include tabulation of the number of events, by type of event (hospitalization, death).

The analysis will be carried out using the following sample SAS code:

```
proc freq data=response_tabulated_by_stratum_and_trt;
   tables strata*trt01a*resp /
          relrisk cmh(mantelfleiss) oddsratio(cl=score(correct=no))
relrisk(cl=score(correct=no))
                alpha=0.05 commonriskdiff(cl=klingenberg test correct=no
column=1);
   weight Count;
run;
```

Each region-age combined strata has a 2×2 table with counts for BEM hospitalized, BEM not hospitalized, Placebo hospitalized, and Placebo not hospitalized. In total there will be 8 strata. When low number of events happens and the model can't be carried out, removal of strata will be carried out in the following order: age first, then region. If all strata are indicated to be removed, then an unstratified comparison will be carried out.

Intercurrent events are those that occur after treatment initiation and either preclude observation of the variable or affect its interpretation. For the intercurrent events of rescue (e.g., intervention with remdesivir; these will be identified by blinded medical review prior to analysis), a treatment policy approach will be taken, whereby use of rescue and concomitant medications will be ignored, and observations collected after they are administered to the subject will be used. In addition, subjects will be followed, assessed, and analyzed irrespective of their completion of the course of the study treatment. The analysis will be performed based on observed data, and the following measures will be used to reduce missing hospitalization/death data:

- eCRFs capturing clinical information that may be indicative of hospitalization will be reviewed to mitigate the risk of unreported hospitalizations
- study staff may use a public information source to obtain information about survival status if a patient withdraws from the study.

8.6.2. Sensitivity Analyses of the Primary Efficacy Results

Planned supplemental analyses to support the primary endpoint result are identified in Table 4 . The supplemental analyses investigate the impact of significant COVID-19 rescue therapy and data handling for early study discontinuations. For the CMH and stratified log-rank analyses, incorporation of strata will be the same as that used for the primary analysis. The Kaplan-Meier and log-rank analyses will include all data to Day 29.

The primary endpoint may also be analyzed for each of the subgroups defined in Section 8.4 .

8.7. Analysis of the Secondary Efficacy Endpoints

The following secondary endpoints will be analyzed using the same approach as that for the primary endpoint:

- Proportion of subjects in the supportive-care-only population with all-cause hospitalization or all-cause death through Day 60
- Proportion of subjects in the supportive-care-only population with all-cause hospitalization, all-cause death, or COVID-19 related medically attended visits through Day 29 and Day 60.
- Proportion of subjects in the supportive-care-only population with COVID-19-related hospitalization or who died due to any cause death through Day 29 and Day 60
- Proportion of subjects in the supportive-care-only population who died due to any cause through Day 29 and Day 60
- Proportion of subjects in the supportive-care-only population with COVID-19-related complications (e.g., death, hospitalization, radiologically confirmed pneumonia, acute respiratory failure, sepsis, coagulopathy, pericarditis/myocarditis, cardiac failure) through Day 29 and Day 60
- Proportion of subjects in the supportive-care-only population with COVID-19-related medically attended visits (hospitalization, ER visit, urgent care or physician’s office visit or telemedicine visit) or who died due to any cause through Day 29 and Day 60.
- Proportion of subjects in the supportive-care-only population with COVID-19-related medically attended visits through Day 29 and Day 60
- Proportion of subjects in the supportive-care-only population with COVID-19 symptom relapse through Day 29
- Proportion of subjects in the supportive-care-only population with viral load rebound through Day 29
- Proportion of subjects in the combination-antiviral population with COVID-19 symptom relapse through Day 29
- Proportion of subjects in the combination-antiviral population with viral load rebound through Day 29

COVID-19 related complications will be adjudicated by blinded medical review prior to analysis.

The proportion of subjects with COVID-19 symptom relapse in each population through Day 29 will be tabulated. This summary will be based on the response to "Have your COVID-19 symptoms returned after previously improving?" (i.e., question 4 in the Symptom Diary, see Section 11.1).

The proportion of subjects with viral load rebound through Day 29 will be tabulated. Viral load rebound is defined as a ≥1 $\log_{10}$ increase in viral load (measured by RT-qPCR) at any non-missing time point (including unscheduled assessments) after EOT. Subjects must have at least one post-EOT measurement result to be evaluated for viral load rebound.

8.8. Analysis of Exploratory Efficacy Endpoints

[redacted]. The following exploratory endpoints will be examined in both supportive care only and combination antiviral arms.

Time to sustained alleviation or improvement of COVID-19 symptoms

Sustained alleviation/improvement of COVID-19 signs/symptoms is defined as the first occurrence of 4 consecutive days when all symptoms recorded as moderate or severe at study entry are recorded as mild or absent and all symptoms scored mild or absent at study entry are scored as absent. The first day of the 4 consecutive-day period corresponds to the day that the event is achieved. The endpoint will be based on targeted symptoms, which are the 13 symptoms with severities ranging from none to severe in Item 3 of the COVID-19 symptom diary. Subjects who do not achieve the endpoint will be censored according to the following rules:

Case	LOD-3	LOD-2	LOD-1	Last Observed Day (LOD)	Censoring time
1	Not Met	Not Met	Not Met	Not Met	LOD+1
2	Not Met	Not Met	Not Met	Met	LOD
3	Not Met	Not Met	Met	Met	LOD-1
4	Not Met	Met	Met	Met	LOD-2

The cumulative distribution function for time from Day 1 to sustained alleviation/improvement of symptoms will be estimated by Kaplan-Meier methodology and plotted by treatment group. The summary display will include the Kaplan-Meier estimates of proportion achieving the endpoint by day. The analysis will also be carried on subgroups of time from onset of symptoms (<=3 days, > 3days).

Sensitivity analysis will be carried using the same methods for 2 selected subsets of symptoms: 1) runny/stuffy nose, sore throat, cough, feeling hot or feverish, and fatigue; 2) runny/stuffy nose, sore throat, shortness of breath, feeling hot or feverish, headache, and chills; 3) all symptoms from Item 3 of the COVID-19 symptoms diary except fatigue.

Additional subgroup analyses on age group (<65 and >=65), and symptom onset days (<=3 days and >3 days) for time to sustained alleviation of all 13 symptoms will be performed.

Time to self-reported return to usual (pre-COVID-19) health

Self-reported return to usual health is defined as the occurrence of 4 consecutive days when the response to "Have you returned to your usual (pre-COVID) health?" in question 3 of the COVID-19 symptom diary indicates return to usual health. Analysis and data handling will be the same as that described for the endpoint of time to alleviation or improvement of COVID-19 symptoms. The analysis will also be carried on subgroups of time from onset of symptoms (<=3 days, > 3days).

Proportion of subjects requiring post randomization rescue therapy

Post randomization rescue therapies will be identified by blinded medical review. The number and proportion of subjects requiring post randomization rescue therapy will be tabulated.

Duration and severity of COVID-19 symptom relapse

For subjects who meet the criteria of COVID-19 symptom relapse, duration of relapse will be computed by counting the number of days that the relapse question (i.e., question 4 of the COVID-19 symptom diary) is answered "yes". Severity will be based on the average response to question 1 of the COVID-19 symptom diary during the days where relapse is indicated. Subjects with the last record indicating relapse will have duration of relapse right censored. Kaplan-Meier analysis will be used to summarize symptom relapse duration by treatment group. Severity of symptom relapse will be summarized descriptively. The analysis will also be carried on subgroups of time from onset of symptoms (<=3 days, > 3days).

Proportion of subjects with symptom relapse and viral rebound

The proportion of subjects who meet both the symptom relapse and viral rebound criteria will be compared between the BEM and placebo arms using the same method as in the primary analysis. The analysis will also be carried on subgroups of time from onset of symptoms (<=3 days, > 3 days) and baseline viral load groups (<median, >=median).

Time to cessation of SARS-CoV-2 viral shedding by RT-qPCR Time to cessation of SARS-CoV-2 viral shedding will be explored using Kaplan-Meier analysis in two ways: 1) time from first detectable observation of SARS-CoV-2 to the time of the first observation with non-detectable SARS-CoV-2; 2) time from first detectable observation of SARS-CoV-2 to the time of first non-detectable observation of SARS-CoV-2 where the non-detectable status is sustained through Day 29. Where a subject does not meet the cessation endpoint, time to cessation will be censored at the time of the last viral RNA assessment. In cases where the result of SARS-CoV-2 testing is reported as inconclusive, the following rules will be applied:

- If inconclusive observation happens before the first non-detectable observation and all follow up observations are non-detectable, time to sustained cessation is time to the first non-detectable observation.
- If inconclusive observation happens between two non-detectable observations, it is treated the same as non-detectable observation.
- If inconclusive observations happen in the last assessments, time to cessation will be censored at the time of last non-inconclusive assessment.

Change from baseline in amount of severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) viral RNA as measured by reverse transcription quantitative polymerase chain reaction (RTqPCR) over time and Proportion of subjects negative for SARS-CoV-2 viral RNA by RT-qPCR over time

SARS-CoV-2 viral RNA (and change from baseline) will be summarized at each scheduled time point by descriptive statistics on $\log_{10}$ viral load and tabulation of the number and percentage of subjects with non-detectable virus. The analysis will also be carried on subgroups of baseline viral load (<median, >=median).

Area under the curve (AUC) in the amount of SARS-CoV-2 viral RNA as measured by RT-qPCR

Area under the $\log_{10}$ viral load curve will be computed, out to each of Days 5, 7 and 14, and summarized using descriptive statistics. AUC will be derived using the trapezoidal rule and consider actual study day (per Section 6.7.2) as time. The analysis will also be carried on subgroups of baseline viral load (<median, >=median).

[redacted]

8.9. Summary of Reasons for Efficacy Non-Evaluability/Exclusion from Efficacy Analyses

Reasons for exclusion from mITT populations will be tabulated. Categories include did not receive study drug and did not have ≥ 1 positive SARS-CoV-2 result through Day 5 (from central testing on NP swab samples).

9. SAFETY AND TOLERABILITY

9.1. Overall Summary of Tolerability

All adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) and will be summarized by treatment group for the safety population. Only treatment-emergent adverse events (TEAEs) will be summarized. Treatment-emergent adverse events are those with start date/time within the treatment period, defined as the time from first study drug administration to Day 29. Events that occur after Day 29 with relationship to study drug recorded as 'related' will also be considered treatment-emergent. If the AE onset date is missing or partial, the conventions described in Section 6.8.3 will be used to determine if the event is treatment-emergent.

Adverse event severity will be graded according to the Division of AIDS (DAIDS) Table for Grading the Severity of Adult and Pediatric Adverse Events, Version 2.1 dated July 2017[6].

An overall summary of TEAEs will be presented and include tabulations (number and %, unless otherwise indicated) for the following:

- Subjects with at least one TEAE, overall and by worst grade
- Total number (number only) of TEAEs
- Deaths
- Subjects who discontinued study due to a TEAE
- Subjects who discontinued treatment due to a TEAE
- Subjects with a treatment-related TEAE resulting in treatment discontinuation
- Subjects with a treatment-emergent serious adverse event (SAE)
- Subjects with a treatment-related SAE
- Subjects with a treatment-related TEAE, overall and by worst grade
- Subjects with a COVID-19 related TEAE
- Subjects with a non-COVID-19 related TEAE

9.2. Adverse Event Preferred Term and Body/Organ System Summary Tables

Adverse events will be coded by System Organ Class (SOC) and preferred term (PT) using the MedDRA. Summaries by SOC and PT will be provided for the following:

- TEAEs
- TEAEs by maximum severity
- Related TEAEs
- TEAEs leading to study discontinuation
- TEAEs leading to treatment discontinuation

- SAEs
- Deaths.

A table of the most common TEAEs (≥ 5% in any treatment arm) will be presented by PT only with display from most common to least common.

9.3. Exposure

Study drug exposure will be characterized by the following:

- Number of doses of study drug received will be tabulated and summarized using descriptive statistics
- Total amount of study drug received will be summarized using descriptive statistics.

9.4. Concomitant and Other Medications

Concomitant medications (CM) are medications that are administered during the treatment-emergent period. If the CM onset date and/or end date is missing or partial, the conventions described in Section 6.8.2 will be used to determine if the medication is concomitant. Medications will be presented based on WHODRUG coding, summarized using number and percentage of subjects, and listed by subject.

9.5. Routine Laboratory Data

Where applicable, the DAIDS Table for Grading the Severity of Adult and Pediatric Adverse Event, Version 2.1 dated July 2017[6] will be used for assigning toxicity grades (0 to 4) to laboratory results for analysis. Grade 0 will be used to designate a value that does not correspond to a grade. Lab parameters that require clinical input for grading will not be assigned grades; any events from these should be captured as adverse events. For laboratory tests with criteria for both increased and decreased levels, analyses for each direction (i.e., increased, decreased) will be presented separately.

The following summaries will be provided by treatment group for each of the scheduled chemistry, hematology, and coagulation laboratory parameters:

- Shift from baseline to abnormal high and to worst overall
- Shift from baseline to abnormal low and to worst overall
- Listing of subjects with at least one post-baseline clinically significant abnormal laboratory observation
- Shift from baseline at each visit and to worst overall toxicity grade (for quantitative parameters with available DAIDS toxicity grades)
- Listing of subjects with at least one post-baseline laboratory observed value meeting a DAIDS toxicity grade ≥ 3 (for quantitative parameters with available DAIDS toxicity grades)

- Maximum post-baseline ALT/AST observed value categorized as < 3 x upper limit of normal (ULN), ≥ 3 to < 5 x ULN, ≥ 5 to < 10 x ULN or ≥ 10 ULN. Maximum post-baseline total bilirubin (TBL) observed value categorized as < 2 x ULN or ≥ 2 x ULN
- Listing of subjects with at least one post-baseline observed ALT value > 3 x ULN, AST value > 3 x ULN or TBL value ≥ 2 x ULN will be provided
- Number and percentage of subjects in each laboratory parameter category by visit (for categorical measurements only).

9.6. Vital Signs

The following summaries will be provided for each scheduled vital sign parameter:

- Observed and change from baseline by visit
- Number and percentages of subjects with at least one markedly abnormal post-baseline observed value/change from baseline (refer to Table 5)
- Listing of subjects with at least one markedly abnormal observed value/change from baseline.

All vital sign data will be listed.

Table 5: Criteria for Markedly Abnormal Changes in Vital Signs

Parameter	Low	High
SBP	≤ 90 mmHg AND change from baseline ≤ -20 mmHg	≥ 180 mmHg AND change from baseline ≥ 20 mmHg
DBP	≤ 50 mmHg AND change from ≤ -15 mmHg	≥ 105 mmHg AND change from baseline ≥ 15 mmHg
Pulse rate	≤ 50 bpm AND change from baseline ≤ -15 bpm	≥ 120 bpm AND change from baseline ≥ 15 bpm
Oxygen saturation	< 93 %	Not applicable
Body temperature	Not applicable	≥ 38.3 °C AND change from baseline ≥ 1.1 °C

9.7. Physical Examination

Physical examination data will be listed.

9.8. Pharmacokinetics

Descriptive summary statistics for concentration by time intervals from the most recent dose will be presented for:

- Plasma concentrations of AT-511 (free base form of AT-527 [BEM]) and its metabolites based on sparse pharmacokinetic (PK) sampling at steady-state**.** The time intervals to be summarized are: 0 – 0.5 hours, >0.5 – 1 hours, >1–2 hours, >2 – 4 hours, >4 – 8 hours, >8 – 12 hours, and >12 hours.

Other analyses of PK data from this study (e.g., population PK analysis) will be described in a separate document, if applicable.

10. REFERENCES

1. International Conference on Harmonization. Statistical Principles for Clinical Trials. Step 5, dated September 2008.
2. [Paxlovid. Prescribing Information] Pfizer; 2023. Accessed June 1, 2023. . https://labeling.pfizer.com/ShowLabeling.aspx?id=19599
3. Lagevrio EUA 2022. US FDA Emergency Use Authorization for Lagevrio, Revised Authorization 23 Mar 2022. https://www.fda.gov/media/157300/download
4. Gottlieb RL, Vaca CE, Paredes R, et al. Early remdesivir to prevent progression to severe Covid-19 in outpatients. N Engl J Med 2022;386:305-15
5. Klingenberg B. A new and improved confidence interval for the Mantel-Haenszel risk difference. Stat Med. 2014;33(17):2968-2983. doi:10.1002/sim.6122
6. U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Allergy and Infectious Diseases, Division of AIDS. Division of AIDS (DAIDS) Table for Grading the Severity of Adult and Pediatric Adverse Events, Version 2.1. Available from: https://rsc.niaid.nih.gov/sites/default/files/daidsgradingcorrectedv21.pdf

11. APPENDICES

11.1. COVID-19 Symptom Diary

Date of completion:_____ Time of completion:_____

SYMPTOMS

1. Please choose the response below that best describes the worst severity of your COVID-19 symptoms over the past 24 hours.

☐ No symptoms	☐ Mild	☐ Moderate	☐ Severe

2. Please choose the response below that best describes your general physical health over the past 24 hours.

☐ Excellent	☐ Very good	☐ Good	☐ Fair	☐ Poor

3. Have you returned to your usual (pre-COVID) health?

☐ Yes	☐ No

Indicate the worst severity of your symptoms **DURING THE PAST 24 HOURS**

Symptoms	**Absent**	**Mild**	**Moderate**	**Severe**
Cough	☐	☐	☐	☐
Shortness of breath or difficulty breathing	☐	☐	☐	☐
Feeling feverish	☐	☐	☐	☐
Chills	☐	☐	☐	☐
Fatigue (low energy)	☐	☐	☐	☐
Body pain or muscle pain or aches	☐	☐	☐	☐
Diarrhea (loose or watery or more frequent stools)	☐	☐	☐	☐
Nausea (feeling like you want to throw up)	☐	☐	☐	☐
Vomiting (throwing up)	☐	☐	☐	☐
Headache	☐	☐	☐	☐
Sore throat	☐	☐	☐	☐
Nasal obstruction or congestion (stuffy nose)	☐	☐	☐	☐
Nasal discharge (runny nose)	☐	☐	☐	☐

Indicate if you have had the following during the past 24 hours:

	Yes	No
Loss of taste	☐	☐
Loss of smell	☐	☐

4. Symptom Relapse: Have your COVID-19 symptoms returned after previously improving?

☐ Yes	☐ No	☐ Not applicable (symptoms have not yet improved)

11.2. Sample Size Calculation Details

The base sample size to have 90% power to detect a statistically significant reduction in the rate of hospitalization/death, given the study assumed rates of hospitalization, based on a one-sided 0.025 level test was computed using proc power is SAS (Version 9.4). This indicated that 991 subjects per arm (i.e. 1982 total subjects with 1:1 randomization). The corresponding code and output were:

```
proc power;
twosamplefreq test=pchi
power=0.9
groupproportions = (0.016, 0.04)
alpha=0.025
sides=1
npergroup=.
;
run;
```

The SAS System

The POWER Procedure
Pearson Chi-square Test for Proportion Difference

Fixed Scenario Elements	
Distribution	Asymptotic normal
Method	Normal approximation
Number of Sides	1
Alpha	0.025
Group 1 Proportion	0.016
Group 2 Proportion	0.04
Nominal Power	0.9
Null Proportion Difference	0

Computed N per Group	
Actual Power	N per Group
0.900	991

The gsDesign function in the "gsDesign" package in R (Version 4.0.3) was used to identify the sample size adjustment and testing criteria needed to implement the planned interim analyses at 30% and 65% with O'Brien-Fleming spending for both alpha and beta (non-binding). The Lan-Demets spending function is used to approximate the O'Brien-Fleming boundaries (specified as sfLDOF in the code below). This indicated that a sample size inflation of 1.055 (at final) is needed for 90% power with overall type 1 error at one-sided 0.025 level. Total beta spending was specified at a 0.10 level. The corresponding code and output were:

```
> gsDesign(k=3, test.type=4, alpha=0.025, beta=0.1, sfu=sfLDOF, sfl=sfLDOF, sflpar=-4, timing=c(0.3, 0.65,1))
Asymmetric two-sided group sequential design with
90 % power and 2.5 % Type I Error.
Upper bound spending computations assume
trial continues if lower bound is crossed.

           Sample
            Size     ----Lower bounds----  ----Upper bounds-----
  Analysis Ratio*   Z   Nominal p Spend+  Z   Nominal p Spend++
         1  0.317 -0.96    0.1681 0.0027 3.93    0.0000  0.0000
         2  0.686  0.94    0.8262 0.0387 2.55    0.0054  0.0054
         3  1.055  1.99    0.9767 0.0587 1.99    0.0233  0.0196
     Total                        0.1000                0.0250
+ lower bound beta spending (under H1):
 Lan-DeMets O'Brien-Fleming approximation spending function with none = 1.
++ alpha spending:
 Lan-DeMets O'Brien-Fleming approximation spending function with none = 1.
* Sample size ratio compared to fixed design with no interim

Boundary crossing probabilities and expected sample size
assume any cross stops the trial

Upper boundary (power or Type I Error)
          Analysis
   Theta      1      2      3  Total   E{N}
  0.0000 0.0000 0.0054 0.0180 0.0234 0.6856
  3.2415 0.0177 0.5368 0.3455 0.9000 0.8277

Lower boundary (futility or Type II Error)
          Analysis
   Theta      1      2      3  Total
  0.0000 0.1681 0.6591 0.1494 0.9766
  3.2415 0.0027 0.0387 0.0587 0.1000
```

With a base sample size of 1982, inflation of 1.055 indicates a required sample size of 2100 evaluable subjects for the primary analysis.