Official Title: A Multicenter, Phase III Randomized, Double-Blind, Placebo-Controlled, Outpatient Study to Evaluate the Efficacy, Safety, Antiviral Activity of RO7496998 (AT-527) in Patients With Mild or Moderate COVID-19

NCT Number: NCT04889040

Document Date: SAP Version 1: 22-June-2021

STATISTICAL ANALYSIS PLAN

STUDY TITLE: **A MULTICENTER, PHASE III RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, OUTPATIENT STUDY TO EVALUATE THE EFFICACY, SAFETY, AND ANTIVIRAL ACTIVITY OF RO7496998 (AT-527) IN PATIENTS WITH MILD OR MODERATE COVID-19**

STUDY NUMBER:	CV43043
STUDY NAME:	MORNINGSKY
VERSION NUMBER:	1
ROCHE COMPOUND:	AT-527 (RO7496998)
EUDRACT NUMBER:	2020-005759-18
IND NUMBER:	149508
NCT NUMBER:	To be determined
PLAN PREPARED BY:	[redacted], M.Sc.

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
22-Jun-2021 15:30:08	Company Signatory	[redacted]
[redacted]	[redacted]	
Date	Co-Approver, [redacted], Atea Pharmaceuticals, Inc.	

SPONSOR:	F. Hoffmann-La Roche Ltd (Ex-United States) Atea Pharmaceuticals, Inc. (United States)
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland 125 Summer Street, Boston, Massachusetts 02110, U.S.A.
DATE FINAL:	See electronic date stamp above

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This SAP was developed based on Roche SAP model document Version 2, 26 October 2020.

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
1	see electronic date stamp on title page	Version 3, 2 June 2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
AE	adverse event
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
CL/F	apparent clearance
V_{ss}/F	apparent steady-state volume of distribution
AUC	area under the concentration curve
AST	aspartate aminotransferase
BLQ	below the limit of quantification
COVID-19	coronavirus disease 2019
Ct	cycle threshold
eCRF	electronic Case Report Form
EQ-5D-5L	EuroQol 5-Dimension 5-Level
HR	hazard ratio
iDMC	independent Data Monitoring Committee
iDMC	independent Data Monitoring Committee
IxRS	interactive voice or web-based response system
IxRS	interactive voice/web-based response system
iSAP	interim Statistical Analysis Plan
ICH	International Council on Harmonization
IMP	investigational medicinal product
LSM	least squares mean
LOQ	limit of quantification
LLoQ	lower limit of quantification
MedDRA	Medical Dictionary for Regulatory Activities
NLME	Nonlinear mixed effects
PGIS	Patient Global Impression of Severity
PRO	patient-reported outcomes
SpO2	peripheral capillary oxygen saturation
PK	Pharmacokinetic
ROW	rest of the world
RT-PCR	reverse transcriptase polymerase chain reaction
RT-qPCR	reverse-transcriptase quantitative polymerase chain reaction
SAE	serious adverse event
SARS-COV-2	severe acute respiratory syndrome coronavirus-2
SE	standard error
SAP	Statistical Analysis Plan

TTAIS	time to alleviation or improvement of signs and symptoms
VAS	Visual Analog Scale
WBC	white blood cells

1. INTRODUCTION

This Statistical Analysis Plan (SAP) provides details of the planned analyses and statistical methods for the efficacy, safety, antiviral activity, and pharmacokinetic (PK) data for Study CV43043 (MORNINGSKY). Table, listing, and figure specifications are contained within a separate document.

This SAP will be updated or a separate interim SAP (iSAP) will be written if an interim analysis for efficacy and/or virology is carried out.

1.1 OBJECTIVES, ENDPOINTS, AND ESTIMANDS

Table 1 Objectives and Corresponding Endpoints

Primary Objective	Corresponding Endpoint
• To evaluate the efficacy of RO7496998 (AT-527) compared with placebo	• The time to alleviation or improvement of COVID-19 symptoms (Items 1–12 of the COVID-19 Symptom Diary) maintained for a duration of 21.5 hours, defined as follows: – For new symptoms: time from randomization to the alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary) – For preexisting symptoms: time from randomization to when a patient's symptoms have been maintained or improved (Note: Improved requires at least a single category improvement from baseline on the COVID-19 Symptom Diary Likert scale.)
Secondary Objectives	**Corresponding Endpoints**
• To evaluate the efficacy of RO7496998 (AT-527) compared with placebo	• The time to alleviation or improvement of COVID-19 symptoms (Items 1–12 of the COVID-19 Symptom Diary) maintained for a duration of 43 hours, defined as follows: – For new symptoms: time from randomization to the alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary) – For preexisting symptoms: time from randomization to when a patient's symptoms have been maintained or improved (Note: Improved requires at least a single category improvement from baseline on the COVID-19 Symptom Diary Likert scale.) • Time to alleviation of COVID-19 symptoms defined as the time from randomization to the point at which the following criterion is met and maintained for at least 21.5 hours: – Score of 0 or 1 for Items 1–12 of the COVID-19 Symptom Diary • Time to alleviation of COVID-19 symptoms defined as the time from randomization to the point at which the following criterion is met and maintained for at least 43 hours: – Score of 0 or 1 for Items 1–12 of the COVID-19 Symptom Diary • Time to one-category improvement of baseline presenting COVID-19 symptoms (Items 1-12 of the COVID-19 Symptom Diary) maintained for a duration of 21.5 hours defined as time from randomization to when the symptoms have improved by at least one category from baseline on the COVID-19 Symptom Diary Likert scale

Table 1 Objectives and Corresponding Endpoints (cont.)

Secondary Objectives	Corresponding Endpoints
• To evaluate the efficacy of RO7496998 (AT-527) compared with placebo (cont.)	• Time to alleviation of individual symptoms, defined as the time from randomization to the point at which the following criterion is met and maintained (for each individual symptom) for at least 21.5 hours: – Score of 0 or 1 for Items 1–14 of the COVID-19 Symptom Diary • Proportion of patients requiring hospitalization for COVID-19 Proportion of patients with ≥ 1 COVID-19 related medically attended visit through to study end (defined as hospitalization, emergency room [ER] visit, urgent care visit, physician's office visit, or telemedicine visit, with the primary reason for the visit being COVID-19)
	• Duration of fever (time to return to afebrile state [temperature ≤ 37.5°C] and remaining so for at least 21.5 hours) (see Protocol Section 6.4.1) • Frequency of COVID-19 related complications (e.g., death, hospitalization, radiologically confirmed pneumonia, acute respiratory failure, sepsis, coagulopathy, pericarditis/myocarditis, cardiac failure) • Proportion of patients with any post-treatment infection • Proportion of patients with all-cause mortality
• To evaluate the antiviral activity of RO7496998 (AT-527) compared with placebo	• Change from baseline in amount of SARS-CoV-2 virus RNA, as measured by RT-qPCR at each timepoint • Time to cessation of SARS-CoV-2 viral shedding, as measured by RT-qPCR • Proportion of patients positive for SARS-CoV-2 virus RNA, as measured by RT-qPCR at specified timepoints • AUC in the amount of SARS-CoV-2 virus RNA, as measured by RT-qPCR
• To evaluate the safety of RO7496998 (AT-527) compared with placebo	• Incidence and severity of adverse events, with severity determined according to NCI CTCAE v5.0 • Incidence of serious adverse events • Change from baseline in vital signs, including SpO_2 • Change from baseline in targeted clinical laboratory test results
• To characterize the PK profile of AT-511 and major metabolites in plasma	• Plasma concentration of AT-511 (the free base form of RO7496998 [AT-527]), AT-551, AT-229, and AT-273 (a surrogate for the intracellular concentration of the active triphosphate metabolite AT-9010) at specified timepoints
Exploratory Objectives	**Corresponding Endpoints**
• To evaluate the antiviral activity of RO7496998 (AT-527) compared with placebo	• Treatment-emergent amino acid substitutions in SARS-CoV-2 viral genes (nsp12 and potentially other genes) • Anti-SARS-CoV-2 antibody status/titer at specified timepoints • Change from baseline in amount of SARS-CoV-2 virus titer at each timepoint • Time to cessation of SARS-CoV-2 viral shedding, as measured by virus titer

Table 1 Objectives and Corresponding Endpoints (cont.)

Exploratory Objectives	Corresponding Endpoints
• To evaluate the antiviral activity of RO7496998 (AT-527) compared with placebo (contd.)	• Proportion of patients with positive SARS-CoV-2 virus titer at specified timepoints • AUC of SARS-CoV-2 virus titer • Drug susceptibility in patients with evaluable virus at specified timepoints
• To identify and/or evaluate biomarkers that are predictive of response to RO7496998 (AT-527) (i.e., predictive biomarkers), are early surrogates of efficacy, are associated with progression to a more severe disease state (i.e., prognostic biomarkers), can provide evidence of RO7496998 (AT-527) activity (i.e., pharmacodynamic biomarkers), or can increase the knowledge and understanding of disease biology and drug safety	• Relationship between biomarkers in blood, nasopharyngeal swab, Nasosorption™, and saliva samples (listed in Protocol Section 4.5.5) and efficacy, safety, PK, or other biomarker endpoints
• To evaluate health status utility, and PGIS scores of patients treated with RO7496998 (AT-527) as compared with placebo	• Patient's global impression of severity of COVID-19 symptoms as assessed through the use of PGIS at Day 14 and Day 29 • Change from baseline in EQ-5D-5L health utility index-based and VAS scores at specified timepoints
• To support content validity of COVID-19 Symptom Diary	• Optional exit interviews at Day 33 (or within 14 days) in a subset of study patients to gather qualitative data to support content validity and inform interpretation of COVID-19 Symptom Diary scores observed during the study

AUC = area under the concentration–time curve; COVID-19 = coronavirus disease 2019; EQ-5D-5L = EuroQol 5-Dimension, 5-Level Questionnaire; NCI CTCAE v5.0 = National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0; PK = pharmacokinetic; PGIS = Patient Global Impression of Severity; RT-qPCR = reverse-transcriptase quantitative polymerase chain reaction; SARS-CoV = severe acute respiratory syndrome coronavirus; SpO_2 = peripheral capillary oxygen saturation; VAS = Visual Analog Scale.

Table 2 Estimands

Endpoint	Variable or Outcome Measurement	Intercurrent Events		Population-Level Summary for the Variable	Handling of Missing Data
		Treatment Withdrawal/ Rescue/ Concomitant Medications	Death		Study Withdrawal/ Lost to Follow-Up
Primary: TTAIS for 21.5 hours	Time from randomization to alleviation or improvement of COVID-19 symptoms maintained for a duration of 21.5 hours, right censored at day 29 if the patient doesn't meet criteria	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.	Hypothetical approach where patient censored at Day 29 if intercurrent event prior to achieving the endpoint	Hazard ratio	Patient censored at Day 29 if study withdrawal is prior to achieving the endpoint
Secondary: TTAIS for 43 hours	Time from randomization to alleviation or improvement of COVID-19 symptoms maintained for a duration of 43 hours, right censored at day 29 if the patient doesn't meet criteria				
Key Secondary: Time to alleviation of symptoms (TTAS) for 21.5 hours	Time from randomization to alleviation of COVID-19 symptoms maintained for a duration of 21.5 hours, right censored at day 29 if the patient doesn't meet criteria				
Secondary: TTAS for 43 hours	Time from randomization to alleviation of COVID-19 symptoms maintained for a duration of 43 hours, right censored at day 29 if the patient doesn't meet criteria				

Table 2 Estimands (cont.)

Endpoint	Variable or Outcome Measurement	Intercurrent Events		Population-Level Summary for the Variable	Handling of Missing Data
		Treatment Withdrawal/ Rescue/ Concomitant Medications	Death		Study Withdrawal/ Lost to Follow-Up
Key Secondary: Time to one-category improvement of symptoms	Time from randomization to one-category improvement of baseline presenting COVID-19 symptoms maintained for a duration of 21.5 hours, right censored at day 29 if the patient doesn't meet criteria	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 (contd.)	Hypothetical approach where patient censored at Day 29 if intercurrent event prior to achieving the endpoint (contd.)	Hazard ratio (contd.)	Patient censored at Day 29 if study withdrawal is prior to achieving the endpoint (contd.)
Secondary: Time to alleviation of individual symptoms	Time from randomization to alleviation an individual COVID-19 symptoms maintained for a duration of 21.5 hours, right censored at day 29 if the patient doesn't meet criteria				
Secondary: Time to return to afebrile state [temperature $\leq$ 37.5C]	Time from randomization to return to afebrile state [temperature $\leq$ 37.5C] maintained for a duration of 21.5 hours for a patient with fever at baseline, right censored at day 29 if the patient doesn't meet criteria				
Secondary: Proportion of patients requiring hospitalization for COVID-19	Patients requiring hospitalization for COVID-19 through to study end (including early termination visits)		Composite strategy where patient considered as having the event if they die	Difference in proportion of patients between treatments	The outcome measure up to point of study withdrawal will be used

Table 2 Estimands (cont.)

Endpoint	Variable or Outcome measurement	Intercurrent Events		Population-Level Summary for the Variable	Handling of Missing Data
		Treatment Withdrawal/ Rescue/ Concomitant Medications	Death		Study Withdrawal/ Lost to Follow-Up
Key Secondary: Proportion of patients with ≥ 1 COVID-19 related medically attended visit	Patients with ≥ 1 COVID-19 related medically attended visit through to study end (including early termination visits)	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 (contd.)	Composite strategy where patient considered as having the event if they die (contd.)	Difference in proportion of patients between treatments (contd.)	The outcome measure up to point of study withdrawal will be used (contd.)
Secondary: Proportion of patients with any post treatment infection	Patients with any post treatment infection through to study end (including early termination visits)				
Key Secondary: Frequency (Proportion) of COVID-19 related complications (including hospitalizations and deaths)	Patients with a medically confirmed COVID-19 related complication through to study end (including early termination visits)		Not an intercurrent event as it is the outcome measure		
Key Secondary: Proportion of patients with all-cause mortality	Patients with all-cause mortality through to study end (including early termination visits)				

Table 2 Estimands (cont.)

Endpoint	Variable or Outcome Measurement	Intercurrent Events		Population-Level Summary for the Variable	Handling of Missing Data
		Treatment Withdrawal/ Rescue/ Concomitant Medications	Death		Study Withdrawal/ Lost to Follow-Up
Secondary (Virology): Change from baseline in amount of SARS-CoV-2 virus RNA	Amount of SARS-CoV-2 virus RNA, as measured by RT-qPCR at Day 3, Day 5, Day 7 and Day 14	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.	While on treatment approach where the outcome measure up to point of intercurrent event will be used	Difference in mean change from baseline between treatments	The outcome measure up to point of study withdrawal will be used
Secondary (Virology): AUC in amount of SARS-CoV-2 virus RNA	Amount of SARS-CoV-2 virus RNA, as measured by RT-qPCR to Day 14			AUC for each treatment	
Key Secondary (Virology): Time to cessation of SARS-CoV-2 viral shedding, as measured by RT-qPCR	Time from randomization to cessation of SARS-CoV-2 viral shedding, right censored at Day 14 if the patient doesn't meet criteria		Hypothetical approach where patient censored at Day 14 if the patient doesn't meet criteria if intercurrent event prior to achieving the endpoint	Hazard ratio	Patient censored at Day 14 if the patient doesn't meet criteria if study withdrawal is prior to achieving the endpoint
Secondary (Virology): Proportion of patients positive for SARS-Cov-2 virus RNA	Patients who are positive for SARS-CoV-2 virus RNA, as measured by RT-qPCR at Day 3, Day 5, Day 7 and Day 14		While on treatment approach where the outcome measure up to point of intercurrent event will be used	Difference in proportion of patients between treatments	The outcome measure up to point of study withdrawal will be used

Table 2 Estimands (cont.)

Endpoint	Variable or Outcome Measurement	Intercurrent Events		Population-Level Summary for the Variable	Handling of Missing Data
		Treatment Withdrawal/ Rescue/ Concomitant Medications	Death		Study Withdrawal/ Lost to Follow-Up
Secondary (Safety): Incidence of adverse events	Patients with an adverse event	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.	Not an intercurrent event as it is the outcome measure	Proportion of patients for each treatment	The outcome measure up to point of study withdrawal will be used (contd.)
Secondary (Safety): Severity of adverse events	Severity of adverse events determined according to NCI CTCAE v5.0			Proportion of patients for each treatment for each severity grade	
Secondary (Safety): Incidence of serious adverse events	Patients with a serious adverse event			Proportion of patients for each treatment	
Secondary (Safety): Change from baseline in vital signs	Vital signs, including SpO2		While on treatment approach where the outcome measure up to point of intercurrent event will be used	Means for each treatment	
Secondary (Safety): Change from baseline in targeted laboratory results	Targeted laboratory results				

Table 2 Estimands (cont.)

Endpoint	Variable or Outcome Measurement	Intercurrent Events		Population-Level Summary for the Variable	Handling of Missing Data
		Treatment Withdrawal/ Rescue/Concomitant Medications	Death		Study Withdrawal/ Lost to Follow-Up
Secondary (PK): Plasma concentration of AT-511, AT-551, AT-229, and AT-273	Plasma concentration of AT-511, AT-551, AT-229, and AT-273	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 (contd.)	While on treatment approach where the outcome measure up to point of intercurrent event will be used (contd.)	Means for each treatment (contd.)	The outcome measure up to point of study withdrawal will be used (contd.)

AUC = area under the concentration–time curve; COVID-19 = coronavirus disease 2019; NCI CTCAE= National Cancer Institute Common Terminology Criteria for Adverse Event; PK=pharmacokinetic; RT-qPCR = reverse-transcriptase quantitative polymerase chain reaction ; SARS-CoV-2 = severe acute respiratory syndrome-corona virus-2; SpO2=peripheral capillary oxygen saturation TTAS =time to alleviation of symptoms TTAIS = time to alleviation or improvement of symptoms.

Target Population: Adult and adolescent patients with mild or moderate COVID-19 defined through inclusion/exclusion criteria.

Treatment: Primary treatment effect comparison will be between the RO7496998 (AT-527) arm and the placebo arm regardless of other therapies.

1.2 STUDY DESIGN

This is a Phase III, placebo-controlled, double-blind, multicenter study to assess the efficacy, safety, antiviral activity, and pharmacokinetics of RO7496998 (AT-527) compared with placebo in non-hospitalized adult and adolescent patients with mild to moderate coronavirus disease 2019 (COVID-19) in the outpatient setting.

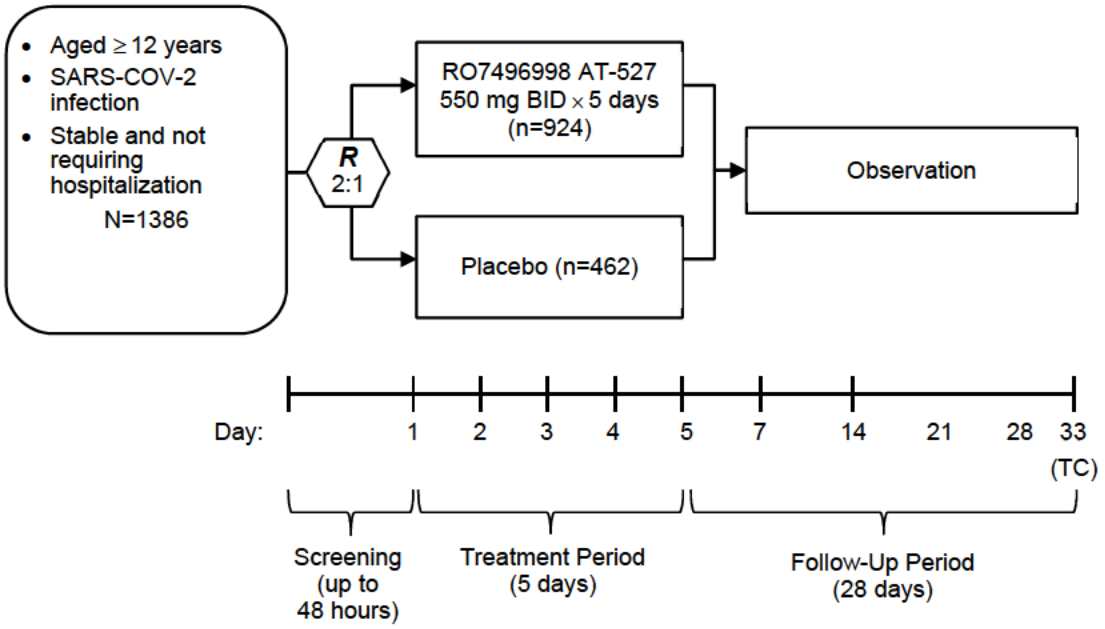
The study will enroll approximately 1386 patients (including approximately 100 adolescents, up to a maximum of 150, with the remaining patients being adults) with confirmed severe acute respiratory syndrome coronavirus (SARS-CoV-2) infection.

The study consists of two periods: a 5-day treatment period and a 28-day safety and efficacy follow-up period, with mandatory study visits on Days 1, 2, 3, 5, 7, and 14. Some of these visits may be conducted by mobile nursing or telemedicine. In addition to mandatory study visits, there will be telephone calls on Day 21 (±2 days) and Day 28 (±2 days). The end of study visit will take place on Day 33 (±3 days), preferentially as a telephone consult but may by conducted in-person in some instances, such as if the patient is enrolling in a longer-term follow-up RO7496998 (AT-527) study. The total study duration for each patient will be approximately 33 days.

The primary analysis will be performed once the data from the last patient's last visit has been collected in the database and all data has been cleaned and verified.

The study schema is shown in Figure 1.

Figure 1 Study Schema



Randomization to 2:1 ratio (active:placebo).

BID = twice a day; N=number of patients; R = randomization; SARS-CoV = severe acute respiratory syndrome coronavirus; TC = telephone call.

1.2.1 Treatment Assignment and Blinding

Patients will be randomly assigned to one of two treatment arms: RO7496998 (AT-527) or matching placebo. Randomization will occur in a 2:1 ratio of RO7496998 (AT-527) to placebo through use of a permuted-block randomization method to ensure a balanced assignment to each treatment arm. Randomization will be stratified by region (North America, Europe, and rest of the world [ROW]) and presence of a high-risk factor (yes, no) for hospitalization due to COVID-19 (see Protocol Sections 3.1 and 3.3.2 for details). A minimum of 40% of the overall study sample will have a high-risk factor for hospitalization because of COVID-19. A minimum of 40% will also be expected for the otherwise healthy population. Stratification caps will be implemented within interactive voice/Web-based response system (IxRS) to ensure these requirements are met.

Study site personnel and patients will be blinded to treatment assignment during the study. The Sponsor and its agents will also be blinded to treatment assignment, with the exception of individuals who require access to patient treatment assignments to fulfill their job roles during a clinical trial (see Protocol Sections 4.2.2 for details).

If unblinding is necessary for a medical emergency (e.g., in the case of a serious adverse event [SAE] for which patient management might be affected by knowledge of

treatment assignment), the investigator will be able to break the treatment code by contacting the IxRS.

As per health authority reporting requirements, the Sponsor's Drug Safety representative will break the treatment code for all serious, unexpected suspected adverse reactions (see Protocol Section 5.7) that are considered by the investigator or Sponsor to be related to an investigational medicinal product (IMP; defined in Protocol Section 4.3).

1.2.2 Sponsor Review and Adjudication

Sponsor representatives from clinical science who are independent of the study team will review data from the study for the purposes of adjudicating adverse events (AEs) identified by the investigator as COVID-19-related complications. This adjudication will determine medically confirmed COVID-19-related complications, which will be used for the frequency of COVID-19-related complications endpoint. These complications and the role and process of Sponsor review and adjudication will be specified in a separate charter.

1.2.3 Data Monitoring

An external independent Data Monitoring Committee (iDMC) will evaluate safety according to policies and procedures detailed in the iDMC charter.

If an interim analysis is conducted for efficacy, the Sponsor will remain blinded. The interim analysis will be conducted by an external statistical group and reviewed by the iDMC. Interactions between the iDMC and Sponsor will be carried out as specified in the iDMC charter.

2. STATISTICAL HYPOTHESES

2.1 PRIMARY ENDPOINT

The primary efficacy analysis for this study will evaluate the time to alleviation or improvement of COVID-19 symptoms maintained for a duration of 21.5 hours of RO7496998 (AT-527) compared with placebo to demonstrate superiority of RO7496998 (AT-527) over placebo.

Let $h_{AT-527}(t)$ be the hazard function for RO7496998 (AT-527) and $h_{PBO}(t)$ be the hazard function for placebo, then the statistical hypothesis tested at a two-sided 0.025 significance level is:

Null hypothesis: The hazard functions for the time to alleviation or improvement of COVID-19 symptoms maintained for a duration of 21.5 hours are the same for RO7496998 (AT-527) arm and placebo arm:

$$H_0: h_{AT-527}(t) = h_{PBO}(t)$$

Alternative hypothesis: The hazard functions for the time to alleviation or improvement of COVID-19 symptoms maintained for a duration of 21.5 hours are not the same for RO7496998 (AT-527) arm and placebo arm:

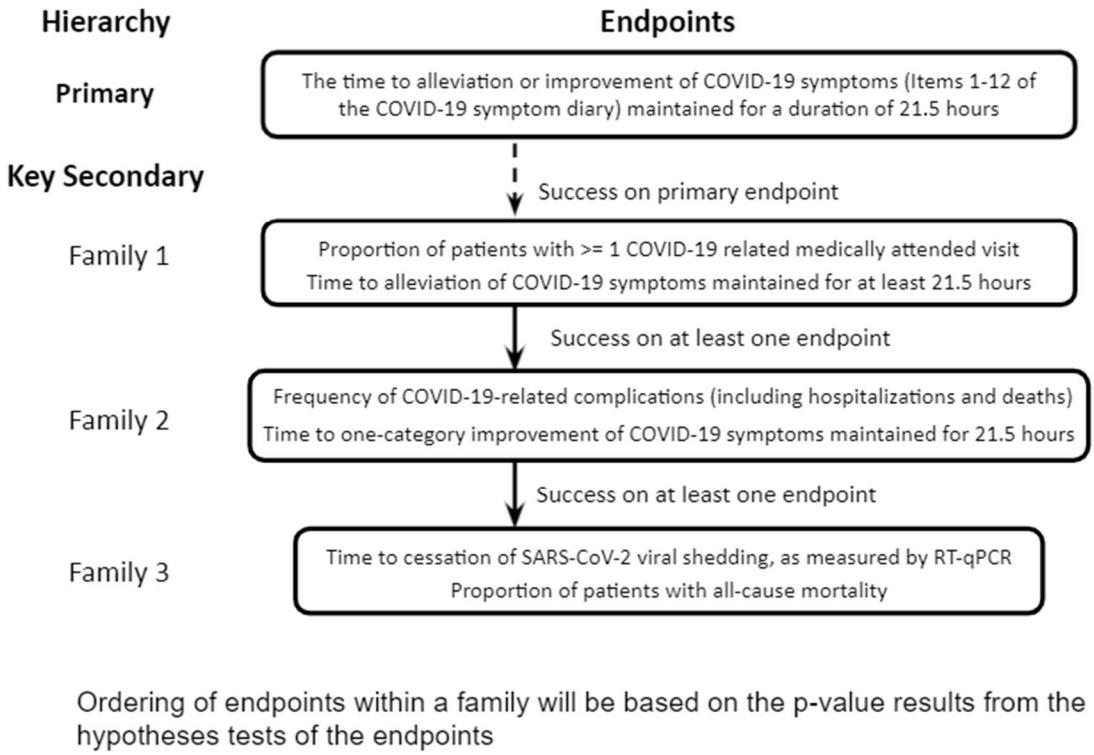
$$H_A: h_{AT-527}(t) \neq h_{PBO}(t)$$

2.2 MULTIPLICITY ADJUSTMENT

The primary and key secondary endpoints (Figure 2) will be evaluated in a hierarchical manner with multiplicity control via multistage gatekeeping under the truncated Holm multiple testing procedure (Dmitrienko et al. 2008). In Families 1 and 2 of key secondary endpoints, the same truncation parameter value of 0.7 will be applied; no truncation is applied in Family 3, since it is prioritized lowest in the hierarchy and imposes no gatekeeping restrictions on subsequent families. The truncation parameter and the relative effect sizes of the endpoints influence how power is balanced over the secondary endpoint families. Truncation parameter values close to 1 should generally be chosen when the effect sizes in high-priority endpoints are relatively large in order to maximize overall power. In other settings where the effect sizes might be smaller or mixed, a truncation parameter value closer to 0 could improve overall power. A parameter value strictly less than 1 in testing families that impose parallel gatekeeping (i.e., at least one endpoint is considered statistically significant after multiplicity adjustment in order to test the next family) is needed to achieve strong overall type I error control (Dmitrienko et al. 2009).

Under this multistage gatekeeping approach, multiplicity-adjusted p-values will reflect the gatekeeping restrictions depicted in Figure 2. In particular, the adjusted p-values for all key secondary endpoints in Families 1, 2, and 3 will be no smaller than the p-value for the primary endpoint. Similarly, the smallest adjusted p-value in Family 2 (3) will be no smaller than the smallest adjusted p-value in Family 1 (2). Endpoints for which the multiplicity-adjusted p-value is greater than 0.025 will not be considered statistically significant. These endpoints and all endpoints not under multiplicity control will be considered to provide supportive information. By repeated application of Proposition 4.1 of Dmitrienko et al. (2008), this testing strategy ensures that the overall type I error is no greater than 2.5%.

Figure 2 Multiple Testing Procedure for Endpoints



COVID-19=coronavirus disease 2019; RT-qPCR=reverse-transcriptase quantitative polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus-2.

3. SAMPLE SIZE DETERMINATION

The required sample size of the efficacy-evaluable set is 1248 patients. It is assumed that the reverse-transcriptase quantitative polymerase chain reaction (RT-qPCR) positive rate for SARS-CoV-2 will be 90%. Therefore, approximately 1386 patients will be randomized to ensure an adequate number of patients in the efficacy-evaluable set. The total number of randomized patients may change based on the percentage of patients who are RT-qPCR positive during the study.

Rationale for the Target Sample Size

The required sample size has been calculated to ensure at least 90% power to detect a hazard ratio of 1.25 in the TTAIS of COVID-19 between the RO7496998 (AT-527) group and the placebo group. This corresponds to a 2-day improvement in the median TTAIS, assuming a median in the placebo group of 10 days (Regeneron 2020). Patients will be randomized to a 2:1 ratio to RO7496998 (AT-527) or placebo, respectively, and the fixed follow-up period will be 29 days. The study will require 1248 patients in the efficacy-evaluable set based on an expected 90% event rate (1122 events required) in order for the log-rank test to have at least 90% power with a two-sided significance level of 0.025, under the assumption of an exponential distribution. Additional patients may be enrolled

in order to observe 1122 events, where an event is meeting the primary endpoint. The minimal detectable difference is a hazard ratio of approximately 1.15, which equates to a difference in medians of approximately 1.33 days under the assumption of a 10-day placebo response.

4. ANALYSIS SETS

The following analysis sets are defined:

Analysis Set	Definition
Efficacy-evaluable	All randomized patients who received at least one dose of study treatment and were centrally assessed as reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) positive for SARS-CoV-2 (by nasopharyngeal swab) at any point during the study, with patients grouped according to randomized treatment.
All patients	All randomized patients, with patients grouped according to randomized treatment.
Safety-evaluable	All randomized patients who received at least one dose of study treatment, with patients grouped to placebo if all doses received were placebo and to AT-527 if at least one dose of AT-527 was received.
PK-evaluable	All randomized patients who received at least one dose of study treatment and have sufficient data to enable estimation of key parameters (e.g., maximum concentration [C_{max}] and minimum concentration [C_{min}]), with patients grouped according to treatment received.

PK=pharmacokinetic; RT-qPCR= reverse transcriptase quantitative polymerase chain reaction; SARS-CoV-2=severe acute respiratory syndrome coronavirus-2.

The efficacy-evaluable set will be used for the primary efficacy analysis. In addition to having been randomized, patients are also required to have received at least one dose of study treatment, so that patients who drop out of the study after randomization but prior to first dose of study treatment will not be included in the analysis. As this is a double-blind study, patients will not know their randomized treatment therefore this cannot influence their decision to take study treatment. Additionally, patients are also required to be centrally assessed as RT-qPCR positive for SARS-CoV-2 (by nasopharyngeal swab) at any point during the study. This cannot be required prior to randomization as this would extend the time between first symptom onset and first dose of study treatment.

In addition, data will be listed for patients who do not receive the treatment they are assigned to at any timepoint.

5. STATISTICAL ANALYSES

Estimands for the primary and secondary endpoints are presented in Table 2.

5.1 GENERAL CONSIDERATION

All statistical tests will be performed at a 0.025 significance level using two-sided tests, unless otherwise noted. As specified in Section 2.2, key secondary endpoints for which the multiplicity-adjusted p-value is greater than 0.025 will not be considered statistically significant. These endpoints and all endpoints not under multiplicity control will be considered to provide supportive information.

Statistical tests will use stratification factors according to IxRS unless there is considerable imbalance in IxRS and electronic Case Report Form (eCRF) stratum classification. 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, continuous variables will be summarized by using the number of non-missing observations, arithmetic mean, standard deviation (SD), median, minimum, 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.

All efficacy and virology analyses will be performed on the efficacy-evaluable set, all patient disposition and study conduct analyses will be performed on the All Patients set, and all safety analyses will be performed on the safety-evaluable set, unless otherwise specified.

Missing data due to study withdrawal or lost to follow up will be handed as described in Table 2. Any other missing data will not be imputed unless otherwise specified in the specific section describing an endpoint.

Time to event secondary efficacy and virology endpoints will use the same statistical methods as used for the primary endpoint as described in Section 5.3.2, and will be measured in days unless otherwise stated.

For time to event efficacy endpoints, once the first instance of the criteria has been met at an assessment, assessments with endpoint data present prior to the assessment that demonstrates the criteria is maintained for the required duration (either ≥21.5 hours or ≥43 hours as appropriate), must also meet the criteria to meet the endpoint i.e., subsequent assessments after a patient first meets the criteria but prior to the assessment showing maintenance of that criteria for 21.5 hours (or 43 hours if appropriate) must also meet the criteria if there is non-missing data at these assessments.

In addition, for time to event efficacy endpoints, once the endpoint has been met, subsequent data for a patient that does not meet the efficacy criteria will not result in a

null endpoint. Due to this, these endpoints can also be thought of as the time to first instance of a maintained defined criteria (such as alleviation) where maintained here refers to the required duration.

Secondary efficacy and virology endpoints that evaluate proportion of patients will be analyzed using a Cochran-Mantel-Haenszel test adjusted by the stratification factors for region (North America, Europe, ROW) and high-risk factor (yes, no), unless stated otherwise. The proportion in each treatment arm together with the difference in proportions, odds ratio and accompanying 97.5% CIs will be presented along with the p-value.

Continuous secondary virology endpoints, such as change from baseline endpoints, will be analyzed using an analysis of covariance (ANCOVA) model with region (North America, Europe, and ROW), high-risk factor (yes, no) as stratification variables, and the baseline value of the studied measure as a covariate. The least squares mean (LSM) change and its standard error (SE) of means in each group, the difference between two groups in least squares mean (LSM change), SE, and 97.5% CI, and p-values will be presented. Additionally, descriptive summaries for the change from baseline will also be presented by treatment group.

Area under the concentration curve (AUC) for secondary and exploratory endpoints will be analyzed using the van Elteren test where AUC is calculated using the trapezoidal method. AUC of change from time 0 (t_0) to time K (t_k) is given by the formula:

$$\sum_{k=1}^{K} \frac{(y_k + y_{k-1})(t_k - t_{k-1})}{2},$$

where t_k (hours) represents the date of the k^{th} virus RNA (reverse-transcriptase polymerase chain reaction [RT-PCR]) assessment ($k = 0, \ldots, K$) and y_k represents the $\log_{10}$ value of the k^{th} virus RNA (RT-PCR) assessment. 24-hours of time will be converted into one day.

Descriptive summary statistics of the AUC will be presented for each treatment group along with the p-value from the van Elteren test.

All secondary virology endpoints will be evaluated on the nasopharyngeal swab samples. Additional descriptive analyses for saliva and serum samples will be performed on an exploratory basis for the secondary virology endpoints, as required.

The qualitative result of the SARS-CoV-2 reverse-transcriptase polymerase chain reaction (RT-PCR) test is determined using cycle threshold (Ct) values. If the qualitative result is positive, then an associated quantitative result is also provided. The quantitative result is either numerical or below the limit of quantification (BLQ). The limit of quantification (LOQ) for the virus RNA is 120 copies/ml. If the virus RNA sample has a

positive qualitative result, but a quantitative result BLQ, virus RNA will be imputed as LOQ – 1 (119 copies/mL). Negative qualitative samples will be imputed as LOQ/2 (60 copies/ml) before being transformed to the log10 scale.

Baseline is defined as the last available assessment prior to first dose of study treatment. For efficacy analyses, if there is no such assessment available then baseline can be up to 4 hours post first dose, with the earliest assessment used as baseline in the case of multiple assessments within this time.

Secondary safety endpoints and other AE analyses will be assessed through descriptive summaries of treatment-emergent adverse events unless otherwise stated. A treatment-emergent adverse event is defined as any new adverse event reported or any worsening of an existing condition on or after the first dose of study treatment. Data reported will include data from baseline up until the patient completes/discontinues from the study and including data from the early termination visit, if applicable. All verbatim adverse event terms will be mapped to Medical Dictionary for Regulatory Activities (MedDRA) Preferred Terms, and adverse event severity will be graded according to the NCI CTCAE v5.0 scale. Adverse events will be tabulated by system organ class (in decreasing order of overall incidence) and by MedDRA Preferred Term (in decreasing order of overall incidence), unless otherwise stated.

Data will be analyzed per the reported assessment; time windows will not be implemented unless otherwise specified.

5.1.1 COVID-19 Symptom Diary

Patients complete the COVID-19 Symptom Diary up to three times on Day 1, twice per day on Days 2–14 and once per day on Days 15–29. Expectation for assessments is as follows:

- Day 1, up to two of the following:
  - Prior to first dose assessment
  - Bedtime assessment
- Days 2–14:
  - Morning assessment
  - Bedtime assessment
- Days 15–29:
  - Bedtime assessment

Data will be analyzed as it was captured per the expected assessments above; time windows will not be implemented.

Missing data for individual symptoms at an assessment where data for other symptoms from the COVID-19 Symptom Diary are present will be set to the highest score on the Likert scale.

Missing data due to a missing assessment (i.e., all symptoms missing) where the assessment would be the first expected assessment to show efficacy criteria maintained for ≥21.5 hours (or ≥43 hours as appropriate) will be assumed to have met the efficacy criteria as long as the data at the next expected assessment following the missed assessment also has met the efficacy criteria.

5.2 PARTICIPANT DISPOSITION

Summaries of patient disposition, including the number and percentage of patients who are randomized, receive treatment, complete treatment, discontinue treatment, complete study, discontinue study, and the associated reasons for discontinuing treatment and study will be created. Summaries of study duration (in days) defined as the date of the study completion/discontinuation minus the date of first study treatment plus one, and visit disposition will also be created. Patient disposition summaries will be created using the efficacy-evaluable and the All Patients set, and will be tabulated by treatment group separately as well as combined.

5.3 PRIMARY ENDPOINT(S) ANALYSIS

The primary efficacy objective of this study is to evaluate the efficacy of RO7496998 (AT-527) compared with placebo in adult and adolescent patients and will be assessed based on the following endpoint:

- The time to alleviation or improvement of COVID-19 symptoms (Items 1–12 of the COVID-19 Symptom Diary) maintained for a duration of 21.5 hours (24 hours minus 10%, to allow some flexibility in the timing of assessments), defined as follows:
  - For new symptoms: time from randomization to the alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary)
  - For pre-existing symptoms: time from randomization to when a patient's symptoms have been maintained or improved (note improved requires at least a single category improvement from baseline on the COVID-19 Symptom Diary Likert scale)

The primary efficacy endpoint definition includes both alleviation of new COVID-19 symptoms and maintenance/improvement of pre-existing COVID-19 symptoms to allow for the possibility that some patients may have concurrent conditions with symptoms similar to those observed with COVID-19 (e.g., cough in a patient with chronic obstructive pulmonary disease). At screening, patients will be assessed with a 14-item COVID-19 symptom severity assessment to identify pre-existing symptoms (within the prior 30 days), and assess if they worsened due to COVID-19. Symptoms that are not pre-existing are considered to be new symptoms and need to achieve sustained alleviation (score of 0 or 1) to meet the endpoint.

If a symptom is pre-existing, the patient is asked if the severity at screening is worse than what was experienced within the last 30 days prior to COVID-19:

- If the severity is not worse, the symptom must be maintained (no worsening) or improved to meet the endpoint.
- If the severity is worse, then a sustained improvement of at least one grade must be achieved to meet the endpoint. For example, a symptom classed as severe at screening must improve to moderate, mild, or none, and a symptom classed as moderate at screening must improve to mild or none.

The hazard ratio for TTAIS for RO7496998 (AT-527) and placebo arms will be tested using the stratified log-rank test with three regions (i.e., North America, Europe, and ROW) and high-risk factor (yes, no) included as the stratification factors. Kaplan-Meier plots, median time to response, 97.5% CIs, and p-values will be presented.

The log-rank test is most powerful when the assumption of proportional hazards holds. The proportional hazards assumption will be tested graphically using the log-cumulative hazard plot by treatment group. Should the proportional hazards assumption not hold, the Gehan-Wilcoxon test will be used to analyze the data.

The primary estimand is defined in Table 2. The primary efficacy analysis population will be the efficacy-evaluable set, which consists of all patients who were randomized to treatment, received a dose of study treatment, and were RT-qPCR positive for SARS-CoV-2 at any point during the study.

5.3.1 Definition of Primary Endpoint

The time to alleviation or improvement of COVID-19 symptoms is defined as the time from randomization to the first time at which all COVID-19 symptoms from Items 1-12 of the COVID-19 Symptom Diary were either alleviated, maintained or improved for a minimum duration of 21.5 hours (24 hours minus 10%, to allow some flexibility in the timing of assessments). Alleviation, maintenance, or improvement of symptoms is defined as the time when all 12 COVID-19 symptoms from Items 1-12 of the COVID-19 Symptom Diary have been assessed by the patient as follows:

- For pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) judged by the patient to be worse at screening:
  - Improvement of at least one grade on the Likert scale from baseline if the severity at baseline is graded moderate or severe:
    - Severe to moderate, mild, or none
    - Moderate to mild or none
    - Mild to mild or none
    - None to mild or none

Note: At screening, patients will only be asked if preexisting symptoms existed (within the last 30 days) and if they were worsened by COVID-19. Patients will be asked to rate the severity at baseline that needs to improve. To avoid recall bias, patients will not be asked to rate the severity of preexisting symptoms prior to COVID-19.

- For pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) judged by the patient NOT to be worse at screening:
  - Maintained (no worsening) baseline severity if the severity at baseline is graded moderate or severe:
    - Severe to severe, moderate, mild, or none
    - Moderate to moderate, mild or none
    - Mild to mild or none
    - None to mild or none
- For new symptoms (i.e., symptoms that develop due to COVID-19 which present prior to or post baseline):
  - Alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary):
    - Severe to mild, or none
    - Moderate to mild, or none
    - Mild to mild or none
    - None to mild or none

Where the baseline assessment for COVID-19 symptom severity is missing, the COVID-19 symptoms for a patient will be treated as new symptoms and therefore must meet the alleviation criteria even if symptoms had been identified as existing prior to developing COVID-19.

The primary endpoint will be tested using a two-sided significance level of 0.025.

5.3.2 Main Analytical Approach for Primary Endpoint(s)

The hazard ratio (HR) for TTAIS for RO7496998 (AT-527) and placebo arms will be tested using the stratified log-rank test with the following stratification factors: Region (North America, Europe, and ROW) and high-risk factor (yes, no), unless otherwise required as detailed in Section 5.1 . Kaplan-Meier plots, median time to response with 97.5% CIs as estimated from the Kaplan-Meier curve, and difference in median time between treatment groups with 97.5% CIs using bootstrap method will be presented. In addition, the treatment groups will be compared descriptively using a Cox proportional hazards model adjusting for the stratification factors and a hazard ratio with a 97.5% CI will be produced.

The log-rank test is most powerful when the assumption of proportional hazards holds. The proportional hazards assumption will be tested graphically using the log-cumulative hazard plot by treatment group, where the plots for each treatment group will be parallel if the proportional hazard assumption holds. Should the proportional hazards assumption not hold, the Gehan-Wilcoxon test stratified by Region and high-risk factor will be used to analyze the data.

The primary estimand is defined in Table 2. Summaries of the proportion of patients experiencing an intercurrent event and the time to the intercurrent event will be created. The WHO Drug Global B3 Format dictionary will be used for concomitant medications including rescue medication.

The statistical hypotheses for the primary endpoint is defined in Section 2.1.

5.3.3 Sensitivity Analyses for Primary Endpoint(s)

To support the primary analyses, the following sensitivity analyses will be conducted:

- To test the sensitivity of the results to the special case of handling missing data, all missed assessments will be assumed not to have met the efficacy criteria.
- To test the sensitivity of the results to the COVID-19 symptoms followed, only baseline presenting symptoms are required to meet the relevant criteria.
- To test the sensitivity of the results to the statistical test performed, the stratified Gehan-Wilcoxon test will be used.
- To test the sensitivity of the results to adjustment by the stratification factors, an unstratified log-rank test will be performed.

5.3.4 Supplementary Analyses for Primary Endpoint(s)

5.3.4.1 Subgroup Analyses for Primary Endpoint(s)

The generalizability of primary endpoint results when comparing RO7496998 (AT-527) to placebo will be investigated by estimating the treatment effect in subgroups. Summaries of primary endpoint by the following subgroups will be provided:

- Stratification factors:
  - Region (North America, Europe, and ROW)
  - Presence of a high risk factor for hospitalization due to COVID-19 (yes, no)
- Key baseline demographics:
  - Age (Adolescent 12–<18 years, Adult <65 years, Adult ≥65 years)
  - Sex (Female, Male)
  - Race/Ethnicity
- Disease characteristics:
  - Time from symptom onset to study treatment in days (≤3, >3)
  - Vaccinated for COVID-19 (Yes, No)

- Baseline viral load (high, low; based on median baseline viral load)
- Seropositive status at baseline (Yes, No)
- Presence of a high-risk factor: Obesity (Yes, No)
- Presence of a high-risk factor: Cardiovascular disease (Yes, No)
- Presence of a high-risk factor: Chronic lung disease (Yes, No)
- Presence of a high-risk factor: Chronic metabolic disease (Yes, No)
- Presence of a high-risk factor: Immunocompromised (Yes, No)
- Presence of a high-risk factor: Chronic liver disease (Yes, No)
- Presence of a high-risk factor: Chronic kidney disease (Yes, No)

5.3.4.2 Supplementary Analyses for Primary Endpoint(s)

To support the primary analyses, the following supplementary analyses will be conducted:

- For intercurrent events of study withdrawal, lost to follow-up, and death where censoring is planned for Day 29, censor at the time associated with the intercurrent event.

5.4 SECONDARY ENDPOINT(S) ANALYSIS(SES)

Estimands for all secondary endpoints are defined in Table 2.

5.4.1 Key/Confirmatory Secondary Endpoint(s)

5.4.1.1 Proportion of Patients with ≥1 COVID-19-Related Medically Attended Visit

COVID-19-related medically attended visits are defined as hospitalization, emergency room (ER) visit, urgent care visit, physician's office visit, or telemedicine visit, with the primary reason for the visit being COVID-19. The statistical method is detailed in Section 5.1.

5.4.1.2 Time to Alleviation of Symptoms Maintained for a Duration of 21.5 Hours

Time to alleviation of symptoms defined as the time from randomization to the point at which a score of 0 or 1 for all Items 1-12 of the COVID-19 Symptom Diary is met and maintained for at least 21.5 hours.

For new symptoms (i.e., symptoms that develop due COVID-19 which present prior to or post baseline) and pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) regardless of whether judged to be worse at screening:

- Alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary):
  - Severe to mild, or none
  - Moderate to mild, or none

- Mild to mild or none
- None to mild or none

The same statistical methods as for the primary endpoint will be used as described in Section 5.3.2.

5.4.1.3 Frequency (Proportion) of Medically Confirmed COVID-19-Related Complications

COVID-19-related complications are defined as deaths and hospitalizations that are reported as COVID-19-related complications by the investigator, and medically confirmed COVID-19-related complications of radiologically confirmed pneumonia, acute respiratory failure, sepsis, coagulopathy, pericarditis/myocarditis, cardiac failure etc. The process of medical confirmation of certain events is briefly described in Section 1.2.2. The statistical method is detailed in Section 5.1. A sensitivity analysis of this endpoint using COVID-19-related complications as reported by the investigator will also be produced.

5.4.1.4 Time to One-Category Improvement of Baseline Presenting Symptoms Maintained for 21.5 Hours

Time to one-category improvement of baseline presenting COVID-19 for Items 1-12 of the COVID-19 Symptom Diary maintained for a duration of 21.5 hours defined as time from randomization to when the symptoms have improved by at least one category from baseline on the Likert scale, where improvement is defined as:

- For pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) judged by the patient to be worse at screening OR for new symptoms (i.e., symptoms that develop due COVID-19 which present prior to or post baseline):
  - Improvement of at least one grade on the Likert scale from baseline if the severity at baseline is graded moderate or severe:
    - Severe to moderate, mild, or none
    - Moderate to mild or none
    - Mild to mild or none
    - None to mild or none
- For pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) judged by the patient NOT to be worse at screening:
  - Maintained (no worsening) baseline severity:
    - Severe to severe, moderate, mild, or none
    - Moderate to moderate, mild or none
    - Mild to mild or none
    - None to none

Where the baseline assessment for COVID-19 symptom severity is missing, the patient will be excluded from the analysis.

The same statistical methods as for the primary endpoint will be used as described in Section 5.3.2.

5.4.1.5 Time to Cessation of SARS-CoV-2 Viral Shedding as Measured by RT-qPCR

Time to cessation of viral shedding by RT-PCR is defined as the time from randomization to the first time when a negative qualitative virus RNA by RT-PCR test result is obtained. The same statistical methods as for the primary endpoint will be used as described in Section 5.3.2.

5.4.1.6 Proportion of Patients with All-Cause Mortality

All-cause mortality is defined as an adverse event captured within the clinical database for which there is a non-missing death date (partial date of death will be considered as a death). The statistical method is detailed in Section 5.1.

5.4.2 Supportive Secondary Efficacy Endpoint(s)

5.4.2.1 Time to Alleviation or Improvement of COVID-19 Symptoms (Items 1–12 of the COVID-19 Symptom Diary) Maintained for a Duration of 43 Hours

The time to alleviation or improvement of COVID-19 symptoms is defined as the time from randomization to the first time at which all COVID-19 symptoms from Items 1-12 of the COVID-19 Symptom Diary were either alleviated, maintained or improved for a duration of 43 hours. Alleviation, maintenance or improvement of symptoms is defined as the time when all 12 COVID-19 symptoms from Items 1-12 of the COVID-19 Symptom Diary have been assessed by the patient as follows:

- For pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) judged by the patient to be worse at screening:
  - Improvement of at least one grade on the Likert scale from baseline if the severity at baseline is graded moderate or severe:
    - Severe to moderate, mild, or none
    - Moderate to mild or none
    - Mild to mild or none
    - None to mild or none

Note: At screening, patients will only be asked if preexisting symptoms existed (within the last 30 days) and if they were worsened by COVID-19. Patients will be asked to rate the severity at baseline that needs to improve. To avoid recall bias, patients will not be asked to rate the severity of preexisting symptoms prior to COVID-19.

- For pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) judged by the patient NOT to be worse at screening:
  - Maintained (no worsening) baseline severity if the severity at baseline is graded moderate or severe:
    - Severe to severe, moderate, mild, or none
    - Moderate to moderate, mild or none
    - Mild to mild or none
    - None to mild or none
- For new symptoms (i.e., symptoms that develop due COVID-19 which present prior to or post baseline):
  - Alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary):
    - Severe to mild, or none
    - Moderate to mild, or none
    - Mild to mild or none
    - None to mild or none

Where the baseline assessment for COVID-19 symptom severity is missing, the COVID-19 symptoms for a patient will be treated as new symptoms and therefore must meet the alleviation criteria even if symptoms had been identified as existing prior to developing COVID-19.

5.4.2.2 Time to alleviation of individual symptoms maintained for a duration of 21.5 hours

Time to alleviation of individual symptoms (Items 1-14 of the COVID-19 Symptom Diary) is defined separately for each symptom as the time from randomization to the point at which a score of 0 or 1 for the individual symptom is met and maintained for at least 21.5 hours in patients with a symptom score $\geq$2 at baseline. The endpoints corresponding to each of the symptoms are as follows:

- Time to alleviation of nasal congestion or runny nose
- Time to alleviation of sore throat
- Time to alleviation of cough
- Time to alleviation of shortness of breath
- Time to alleviation of muscle or body aches
- Time to alleviation of fatigue
- Time to alleviation of headache
- Time to alleviation of chills/sweats
- Time to alleviation of feeling hot or feverish

- Time to alleviation of nausea
- Time to alleviation of vomiting
- Time to alleviation of diarrhea
- Time to alleviation of sense of smell
- Time to alleviation of sense of taste

The same statistical methods as for the primary endpoint will be used as described in Section 5.3.2.

Where a baseline assessment for COVID-19 symptom severity is missing for an individual symptom, the patient will be excluded from the analysis.

5.4.2.3 Time to Alleviation of Symptoms Maintained for a Duration of 43 Hours

Time to alleviation of symptoms is defined as the time from randomization to the point at which a score of 0 or 1 for all Items 1-12 of the COVID-19 Symptom Diary is met and maintained for at least 43 hours.

For new symptoms (i.e., symptoms that develop due COVID-19 which present prior to or post baseline) and pre-existing symptoms (i.e., symptoms that existed prior to developing COVID-19) regardless of whether judged to be worse at screening:

- Alleviation of COVID-19 symptoms (i.e., a score of 0 [none] or 1 [mild] on the COVID-19 Symptom Diary):
  - Severe to mild, or none
  - Moderate to mild, or none
  - Mild to mild or none
  - None to mild or none

The same statistical methods as for the primary endpoint will be used as described in Section 5.3.2.

5.4.2.4 Proportion of Patients Requiring Hospitalization for COVID-19

Hospitalizations for COVID-19 are defined as SAEs for which the investigator has cited that the suspected cause was the disease under study and where there is a non-missing hospital admission date. The statistical method is detailed in Section 5.1.

5.4.2.5 Time to Return to Afebrile State Maintained for at least 21.5 Hours (Duration of Fever)

Time to return to afebrile state (temperature ≤37.5°C) is defined as the time from randomization to the time of afebrile state maintained for at least 21.5 hours in patients with a fever at baseline. The same statistical methods as for the primary endpoint will be used as described in Section 5.3.2.

Where a baseline temperature assessment is missing, the patient will be excluded from the analysis.

5.4.2.6 Proportion of Patients with Any Post-Treatment Infection

Post-treatment infections are defined as any treatment-emergent adverse event with a primary system organ class of infections and infestations. The statistical method is detailed in Section 5.1.

5.4.3 Supportive Secondary Virology Endpoints

5.4.3.1 Change from Baseline in Amount of SARS-CoV-2 Virus RNA as Measured by RT-qPCR at Each Timepoint

The change from baseline in amount of SARS-CoV 2-virus RNA as measured by RT-qPCR will be analyzed at Day 3, Day 5, Day 7 and Day 14 separately. The statistical method is detailed in Section 5.1.

5.4.3.2 Proportion of Patients Positive for SARS-Cov-2 Virus RNA by RT-qPCR at Specified Timepoints

Proportion of patients positive for SARS-CoV-2 virus RNA is defined as the proportion of patients with a positive qualitative virus RNA by RT-PCR test result, which will be analyzed at Day 3, Day 5, Day 7 and Day 14 separately. The statistical method is detailed in Section 5.1.

5.4.3.3 Area Under the Concentration Curve (AUC) in the Amount of SARS-CoV-2 Virus RNA as Measured by RT-qPCR

The statistical method is detailed in Section 5.1.

5.4.4 Supportive Secondary Safety Endpoints

5.4.4.1 Incidence of Adverse Events

Descriptive summaries of treatment-emergent adverse events will be presented.

5.4.4.2 Severity of Adverse Events

Descriptive summaries of treatment-emergent adverse events tabulated by severity will be presented.

5.4.4.3 Change from Baseline in Vital Signs

For each of the vital signs (pulse rate, respiratory rate, systolic and diastolic blood pressure, peripheral capillary oxygen saturation [SpO2], and temperature), descriptive statistics on absolute value and the change from baseline will be presented over time.

For vital signs, including oxygen saturation, time windows for unscheduled and early termination visits are given in Table 3 below. Where multiple measurements within a visit window are available for a patient, the value obtained closest to the target timepoint will be used.

Table 3 Vital Signs (including SpO2) Time Windows

Scheduled Timepoint	Acceptable Time Window
Baseline (Day 1 pre-dose)	Day 1 pre-dose
Day 3	Day 1 post-dose – Day 3
Day 5	Day 4 – Day 5
Day 7	Day 6 – Day 10
Day 14	Day 11 – Day 18

5.4.4.4 Change from Baseline in Targeted Laboratory Results

Laboratory data will use ranges from local laboratories and laboratory values will be converted to Système International units.

Summary tables will present the observed and change from baseline values of the laboratory chemistry parameters: alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, total bilirubin, blood urea nitrogen, creatine kinase, lactate dehydrogenase, C-reactive protein and creatinine, and laboratory hematology parameters: hemoglobin, white blood cell count, and platelets, and laboratory coagulation parameter: D-dimer for baseline and Day 5 visits by treatment group.

For laboratory results, time windows for unscheduled and early termination visits are given in Table 4 below. Where multiple measurements within a visit window are available for a patient, the value obtained closest to the target timepoint (center of the acceptable time window) will be used.

Table 4 Laboratory Time Windows

Scheduled Timepoint	Acceptable Time Window
Baseline (Day 1 pre-dose)	Day 1 pre-dose
Day 5	Day 4 – Day 6

5.4.5 Supportive Secondary PK Endpoints

Samples taken for PK analysis will be processed to plasma and concentrations of AT-511 (free base form of AT-527) and its major metabolites AT-551, AT-229, and AT-273, will be quantified by a validated liquid chromatography-tandem mass spectrometry method.

Descriptive summary statistics for PK concentration data will be provided by visit/sampling time point. Summary statistics will also include geometric mean and coefficient of variation (CV; arithmetic and geometric). Concentrations below lower limit of quantification (LLOQ) will be treated as zero in summary statistics and for PK

parameter calculations. Individual and mean plasma concentration versus time data for AT-511, AT-551, AT-229, and AT-273 will be summarized and plotted.

5.5 EXPLORATORY ENDPOINT(S) ANALYSIS

5.5.1 Exploratory Efficacy Endpoints

- Proportion of patients with ≥1 COVID-19-related medically attended visit (handling missing data due to study withdrawal or loss of follow up with multiple imputation)
- Proportion of patients requiring hospitalizations for COVID-19 (handling missing data due to study withdrawal or loss of follow up with multiple imputation)
- Frequency (Proportion) of medically confirmed COVID-19-related complications (handling missing data due to study withdrawal or loss of follow up with multiple imputation)

5.5.2 Exploratory Virology Endpoints

The exploratory virology endpoints for this study are as follows:

- Treatment-emergent amino acid substitutions in SARS-CoV-2 viral genes (such as *nsp12* and potentially other genes)
- Anti-SARS-CoV-2 antibody status/titer for neutralizing antibodies and anti-S antibodies at specified timepoints
- Change from baseline in amount of SARS-CoV-2 virus titer at each timepoint
- Time to cessation of SARS-CoV-2 viral shedding as measured by virus titer
- Proportion of patients with positive SARS-CoV-2 virus titer at specified timepoints
- AUC of SARS-CoV-2 virus titer
- Drug susceptibility in patients with evaluable virus at specified timepoints

Time to event, incidence, proportion and change from baseline endpoints will be analyzed using the same approach as described for secondary virology endpoints in Section 5.1.

Additional analyses adjusting by baseline amount of virus RNA may also be conducted. Other virology endpoints will be summarized descriptively.

5.5.3 Exploratory Patient Reported Outcome and Health Utility Endpoints

Data from the PGIS and EQ-5D-5L will be used in separate analysis to inform understanding of the COVID-19 Symptom Diary and health utilities. The patient reported outcomes and health utility endpoints for this study are as follows:

- Patient's global impression of severity (PGIS) of COVID-19 symptoms as assessed through the use of the PGIS at Day 14 and Day 29
- Change from baseline in EQ-5D-5L health utility index-based and Visual Analog Scale (VAS) scores at specified timepoints

The EuroQol 5-Dimension 5-Level (EQ-5D-5L) will be scored according to its manual, and results will be reported separately from the clinical study report.

Patient-reported outcomes (PRO) and health utility endpoints will be summarized descriptively.

5.6 OTHER SAFETY ANALYSES

5.6.1 Extent of Exposure

The number of administrations and the duration of exposure to study treatment will be summarized using descriptive statistics. The duration of exposure will be calculated in days as the date and time of last dose of study treatment minus the date and time of first dose of study treatment plus 12 hours and divided by 24 to convert to days.

5.6.2 Adverse Events

Separate summaries will be generated for SAEs, deaths, adverse events leading to discontinuation of study treatment, adverse events leading to discontinuation of study, adverse events of special interest, SAEs suspected to be caused by study treatment and adverse events suspected to be caused by study treatment, as assessed by the investigator. In addition, adverse events and SAEs by time of onset according to the following categories will be reported: Days 1-7, Days 8-14, and Days $\geq$15. If a low number of events is observed, a listing may be produced in place of a summary table. Adverse events will also be listed.

Adverse events of special interest for this study will be identified by the relevant checkbox on the eCRF and are as follows:

- Cases of potential drug induced liver injury that include an elevated alanine aminotransferase (ALT) or aspartate aminotransferase (AST) in combination with either an elevated bilirubin or clinical jaundice, as defined by Hy's Law (see Protocol Section 5.3.5.6)
- Suspected transmission of an infectious agent by the study drug, as defined below
    - Any organism, virus, or infectious particle (e.g., prion protein transmitting transmissible spongiform encephalopathy), pathogenic or non-pathogenic, is considered an infectious agent. A transmission of an infectious agent may be suspected from clinical symptoms or laboratory findings that indicate an infection in a patient exposed to a medicinal product. This term applies only when a contamination of the study drug is suspected.

Additional safety analyses for adverse event and SAEs divided by age and vaccination status subgroup will also be produced.

5.6.3 Laboratory Data

Summaries of the number of patients by NCI CTCAE v5.0 grade for hematology (white blood cells [WBC], hemoglobin, platelets), and for chemistry (alkaline phosphatase, ALT,

AST, total bilirubin, creatinine), and renal function will be produced. The number of patients will be summarized by NCI CTCAE grade category for baseline and worst post baseline value.

Patients with values outside the reference will be listed, with an indication of the direction of the abnormality (high, low).

A listing of patients who meet the pre-specified criteria shown in Table 5 during the study will be presented for each treatment group.

Table 5 Pre-specified Treatment-Emergent Abnormal Laboratory Criteria

AST (U/L) or ALT (U/L) and Total bilirubin (mg/dL)
AST (U/L) or ALT (U/L) > 3 × Baseline
AST (U/L) or ALT (U/L) > 5 × Baseline
AST (U/L) or ALT (U/L) > 3 × Baseline in combination with total bilirubin (mg/dL) > 2 · ULN

ALT= alanine aminotransferase; AST= aspartate aminotransferase; ULN= upper limit of normal.

5.6.4 Vital Signs

A summary of the proportion of patients that have SpO2 ≤ 93% at any point in the study will be created.

5.6.5 Medical History

Medical history data will be summarized descriptively by treatment group using the safety-evaluable set.

5.6.6 Previous and Concomitant Medications

Previous and concomitant treatments will be summarized descriptively by treatment group for the safety-evaluable set. Previous treatments that have been stopped prior to Study Day 1 will be summarized separately, where Study Day 1 is the day of randomization. There will be a summary of all concomitant treatments, including those that were initiated prior to Study Day 1. In addition, there will be summaries of previous and concomitant treatments with the indication given as 'COVID-19' as well as summary of the proportion of patients requiring supplemental oxygen.

5.7 OTHER ANALYSES

5.7.1 Summaries of Conduct of Study

Summary of analysis sets including numbers of patients in each set and a summary of major protocol deviations will be created on the All Patients set. Summaries will be tabulated by treatment group separately as well as combined.

5.7.2 Summaries of Treatment Group Comparability

To review treatment group comparability within the study, a number of variables collected at baseline will be compared across treatment groups on the efficacy-evaluable set.

5.7.2.1 Demographics and Baseline Characteristics

- Age
- Sex
- Race
- Ethnicity
- Region
- Smoking history
- Weight
- BMI

5.7.2.2 Baseline Disease Characteristics

- SARS-CoV-2 RT-PCR Status at Baseline
- Baseline viral load
- Seropositive status at baseline
- Presence of a high-risk factor
- Presence of individual high-risk factors
- Time from symptom onset to study treatment in days
- Number of symptoms present at baseline
- Number of symptoms from pre-existing conditions at baseline
- Number of symptoms with score ≥ 2 present at baseline
- Vaccinated for COVID-19
- Baseline systemic corticosteroid use
- Baseline inhaled corticosteroid use

5.7.3 Pharmacokinetic Analyses

The following pharmacokinetic parameters will be determined on Day 1: maximum concentration (C_{max}), time to maximum concentration (T_{max}) after the first dose from the plasma concentration-time data. At steady state on Day 5 the following PK parameters will be determined: $C_{max,ss}$, $T_{max,ss}$, and $C_{trough,ss}$. All PK analysis will be performed using the actual recorded sampling times for AT-511, AT-551, AT-229, and AT-273. Descriptive summary statistics for PK parameters will include mean (arithmetic and geometric), SD, CV (arithmetic and geometric), median, minimum, and maximum. An exception to this is T_{max} where median, minimum, and maximum will be presented. Inter-patient variability and drug accumulation will be evaluated.

Nonlinear mixed effects (NLME) modeling will be used to analyze the dose-concentration-time data of AT-511 and AT-273 in plasma. Population and individual PK parameters (e.g., apparent clearance [CL/F] and apparent steady-state volume of distribution [V_{ss}/F]) will be estimated and the influence of various covariates (e.g., age, sex, and weight) on these parameters will be investigated. The data collected during this study may be pooled with data collected in other clinical studies (previous Phase I and/or Phase II study in healthy volunteers and patients infected with hepatitis C virus [HCV], COVID-19) as appropriate, to build the population PK model. Secondary PK parameters may be derived from the individual post-hoc predictions. NLME modeling results will be reported in a standalone document distinct from the clinical study report.

Additional PK analyses may be conducted as appropriate. The PK data and parameters derived from these analyses may be used for exploratory graphical analyses of the PD data and parameters.

5.8 INTERIM ANALYSES

5.8.1 Planned Interim Analyses

No interim analysis is planned for this study.

5.8.2 Optional Interim Analyses

To adapt to information that may emerge during the course of this study, the Sponsor may choose to conduct one interim efficacy analysis and/or one interim virology analysis. Full details will be pre-specified in a SAP (either this SAP or a separate iSAP) prior to the interim analysis being conducted and will include rationale and specifications for ensuring the study maintains the highest standards of integrity.

Below are the specifications in place to ensure the study continues to meet the highest standards of integrity when an optional interim analysis is executed.

If an interim analysis is conducted for efficacy, the Sponsor will remain blinded. The interim analysis will be conducted by an external statistical group and reviewed by the iDMC. Interactions between the iDMC and Sponsor will be carried out as specified in the iDMC charter.

The decision to conduct the optional interim efficacy analyses, along with the rationale, timing, and statistical details for the analysis, will be documented in this SAP or in a separate iSAP, and the relevant SAP will be submitted to relevant health authorities prior to the conduct of the interim analysis. The iDMC charter will be updated to document potential recommendations the iDMC can make to the Sponsor as a result of the analysis (e.g., stop the study for positive efficacy, stop the study for futility), and the iDMC charter will also be made available to relevant health authorities.

If there is a potential for the study to be stopped for positive efficacy as a result of the efficacy interim analysis, the type I error rate will be controlled to ensure statistical

validity is maintained. Specifically, the Lan-DeMets α -spending function that approximates the O'Brien-Fleming boundary will be applied to determine the critical value for stopping for positive efficacy at the interim analysis (DeMets and Lan 1994). Additional criteria for recommending that the study be stopped for positive efficacy may be added to the iDMC charter. If the study continues beyond the interim analysis, the critical value at the final analysis would be adjusted accordingly to maintain the protocol-specified overall type I error rate, per standard Lan-DeMets methodology.

If there is a potential for the study to be stopped for futility as a result of the interim analysis, the threshold for declaring futility will include an assessment of the predictive probability that the specified endpoint will achieve statistical significance.

If an interim analysis is conducted for virology, the Sponsor will be unblinded. The interim analysis will be conducted by the Sponsor to confirm adequate antiviral effect. Patients that contribute to the virology interim analysis will be excluded from the final efficacy and virology analyses, and will be replaced in order to maintain power of the primary endpoint analysis and there will therefore be no alpha adjustment for the final analysis. The interim virology analysis will involve an early look at virology parameters over time and may include formal statistical analysis of the change from baseline in viral titer. Details will be specified in a separate iSAP, or by updating this SAP which will be finalized prior to the interim analysis taking place.

6. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

Appendix 1 Changes To Protocol-Planned Analyses

The modified intent-to-treat infected (mITTi) population described in the protocol has been renamed to the efficacy-evaluable set to more accurately describe this analysis set.

The rationale for the target sample size and the primary efficacy endpoint has been updated to remove references to comparing the median TTAIS and to include references to the comparison of the hazard ratio to more accurately describe the comparison to be made and from which the sample size was based upon.

The optional interim analysis was updated to specify that if an optional interim efficacy or virology analysis is conducted that this will be updated in this SAP or within a separate iSAP, which will be finalized prior to the interim analysis taking place.

Appendix 2 COVID-19 Symptom Diary

Do not reproduce or distribute. The Sponsor will provide sites with all instruments to be completed in this study.

Note that the following is a sample. Sites will be provided with separate versions of the diary, reflecting appropriate recall periods.

COVID-19 Symptom Diary (Morning)

This diary will keep track of your COVID-19 symptoms and temperature during the study. Please complete this diary when you first wake up in the morning (before getting out of bed). Remember to take your temperature before you take any medicines to help your COVID-19 symptoms.

Please rate the severity of each symptom at its worst since you went to bed last night.

	None 0	**Mild 1**	**Moderate 2**	**Severe 3**
1. Nasal Congestion or runny nose				
2. Sore Throat				
3. Cough				
4. Shortness of breath (difficulty breathing)				
5. Muscle or body aches				
6. Fatigue (tiredness)				
7. Headache				
8. Chills/Sweats				
9. Feeling hot or feverish				
10. Nausea (wanting to throw up)				
11. Vomiting (thrown up)				
12. Diarrhea (mostly or completely liquid bowel movements)				

Appendix 2 COVID-19 Symptom Diary (Cont.)

	Same as usual 0	Less than usual 1	No sense 2
13. Rate your sense of **smell** since you went to bed last night.			
14. Rate your sense of **taste** since you went to bed last night.			

Appendix 3 Patient Global Impression of Severity

Do not reproduce or distribute. The Sponsor will provide sites with all instruments to be completed in this study.

Patient Global Impression of Severity (Bedtime)

Please rate your overall COVID-19 symptoms at their worst in the past 24 hours.

☐ None

☐ Mild

☐ Moderate

☐ Severe

7. REFERENCES

DeMets DL, Lan KKG. Interim analysis: the alpha spending function approach. Stat Med 1994;13:1341–52.

Dmitrienko A, Tamhane AC, Wiens BL. General multistage gatekeeping procedures. Biom J. 2008;50(5):667–77.

Dmitrienko A, Tamhane AC, Bretz F, editors. Multiple testing problems in pharmaceutical statistics. CRC Press 2009;165–91.

Regeneron Pharmaceuticals, Inc. “REGN-COV2 antibody cocktail program update” Regeneron COVID-19 Program Update. 28 September 2020. Accessed: 16 November 2020. Available from: https://investor.regeneron.com/static-files/a596a85e-e72d-4529-8eb5-d52d87a99070.