Official Title of Study:

Phase 2, Multicenter, Randomized, Double-blind, Placebo-controlled, Phase 2 Study to Investigate the Efficacy and Safety of Asapiprant in Hospitalized Adults With COVID-19

ClinicalTrials.gov Identifier: NCT04705597

Document, Version [Date]: Statistical Analysis Plan, Version 5.0 [01 June 2022]

[redacted]

BIOAGE

Protocol No: BGE-175-201

CP Project ID: BGE20175

A Multicenter, Randomized, Double-blind, Placebo-controlled, Phase 2 Study to Investigate the Efficacy and Safety of Asapiprant in Hospitalized Adults With COVID-19

Statistical Analysis Plan

Version: v5.0

Date: 01JUN2022

SIGNATURE PAGE

Sponsor [redacted]

Signature Date

Biostatistician [redacted]

Signature Date

ABBREVIATIONS

AE	Adverse event
CD	Cluster of differentiation
CI	Confidence interval
C_{max}	Maximum concentration occurring at t_{max}
COVID-19	Coronavirus disease of 2019
CP	Conditional power
eCRF	Electronic case report form
CRO	Contract research organization
CRP	C-reactive protein
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DMC	Data Monitoring Committee
ECG	Electrocardiogram
ECMO	Extracorporeal membrane oxygenation
EMA	European Medicines Agency
EOT	End of treatment
FiO_2	Fractional inspired oxygen
GCP	Good Clinical Practice
ICF	Informed consent form
ICH	International Conference on Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICU	Intensive care unit
IFN-α	Interferon alpha
IFN-γ	Interferon gamma
IL	Interleukin
IMP	investigational medicinal product
IP-10	Interferon gamma-induced protein 10
ITT	Intention to Treat
MCP-1	Monocyte chemoattractant protein-1
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat population
MMRM	mixed model for repeated measures
NG	Nasogastric tube
O_2	Oxygen

PaO_2	Arterial oxygen partial pressure
PD	Pharmacodynamic(s)
PGD2	Prostaglandin D2
PK	Pharmacokinetic(s)
PO	Oral administration
PPS	Per-protocol Set
PT	Prothrombin time
QTcF	QT interval corrected for heart rate by Fridericia's formula
RF	Respiratory failure
SAE	Serious adverse event
SAF	Safety Analysis Set
SAP	Statistical analysis plan
SD	Standard deviation
SOC	Standard of care
SpO_2	Oxygen saturation
SUSAR	Suspected Unexpected Serious Adverse Reaction
TEAE	Treatment-emergent adverse event
TNF-α	Tumor necrosis factor alfa
TSH	Thyroid-stimulating hormone
ULN	Upper limit of normal
WHO	World Health Organization

Table of Contents

MODIFICATION HISTORY

Unique Identifier for this Version	Date of the Document Version	Author	Significant Changes from Previous Authorized Version
1.0	31Mar2021	[redacted]	N/A – First Version
2.0	23Jun2021	[redacted]	Updated per Protocol V4.0 (19May2021)
3.0	09SEP2021	[redacted]	Updated per Protocol V5.0 (07SEP2021)
4.0	14MAR2022	[redacted]	Updated per FDA comments dated November 24, 2021 and BioAge response.
5.0	01JUN2022	[redacted]	1. Update on primary endpoint analysis by removing baseline WHO ordinal scale/ Addendum is added for the justification. 2. Other minor corrections

1 Introduction

This Statistical Analysis Plan (SAP) contains a more technical and detailed elaboration of the principal features of the analysis described in the protocol and includes detailed procedures for executing the statistical analysis. This SAP should be used in conjunction with the protocol. If there are any discrepancies between the protocol and SAP, this SAP will prevail. Any deviations from this SAP that are implemented in the final analysis will be documented with sound clinical and statistical rationale in the Clinical Study Report (CSR). In the event of any changes to the primary endpoints or analyses, these changes will be documented through a protocol amendment, consistent with ICH E9 Statistical Principles for Clinical Trials.

The SAP is based on protocol v5.0, Amendment 04 dated 07 Sep 2021.

Pharmacokinetics (PK) and pharmacodynamics (PD) analyses and reporting are not included in this analysis plan.

A separate document contains the table, figure and listing specifications and any example programming codes.

2 Study Design

This is a randomized, placebo-controlled, parallel-group, multicenter, double-blind study of asapiprant (BGE-175) administered by oral administration (PO) or through a nasogastric tube (NG) in patients ≥ 50 years of age and hospitalized with documented coronavirus disease 2019 (COVID-19) infection who are not yet in respiratory failure (RF).

After signing informed consent, patients will be screened upon presentation at the hospital. Screening will include full physical examination, vital signs, safety laboratory evaluation, oxygen (O2) saturation, pre-diagnostics to measure prostaglandin D2 (PGD2) status, and baseline assessment of World Health Organization (WHO) Ordinal Scale for COVID-19 that is derived from O_2 supplementation requirements. If confirmed that they qualify for this protocol according to listed inclusion and exclusion criteria, patients will receive their first dose of investigational medicinal product (IMP). They will then receive IMP PO or NG (if intubated or unable to swallow medication) once daily, at approximately the same time each day for up to 13 additional days. IMP will be administered in addition to standard of care deemed appropriate by the treating physician(s). Patients will be randomized to receive asapiprant or placebo in a blinded manner. Patients will be monitored daily for all relevant efficacy outcomes, O_2 saturation and adverse events (AEs). Blood will be drawn periodically for safety laboratory measurements, plasma kinetics, lymphocyte subsets, CRP, and cytokines (see protocol, Table 1 Schedule of Assessments). Nasopharyngeal swabs will be collected to measure viral load. Patients will be monitored at follow-up visits on Day 28 (approximately 14 days after the last dose) and Day 57 (approximately 43 days after last dose).

Treatment Period

Each patient will participate in the study for approximately 8 weeks from the time the patient signs the informed consent form (ICF) through the final contact. After screening, each patient will receive, daily, 1 dose of assigned treatment between Day 1 and Day 14. The WHO Ordinal Scale for COVID-19 will be derived daily while hospitalized, at Day 14, and at follow-up visits until the patient has completed the study.

2.1 The Study Objectives

Primary Objective

- To evaluate the efficacy of asapiprant given by PO or NG administration versus placebo on the clinical signs and symptoms of COVID-19 infection in hospitalized patients ⩾ 50 years of age.

Secondary Objectives

- To evaluate the ability of asapiprant given by PO or NG administration to accelerate the clearance of the COVID-19 virus in patients with symptomatic disease requiring hospitalization.
- To determine the efficacy of asapiprant by demonstrating improvement in COVID-19 illness and prevention of progression to RF using various endpoints.
- To evaluate the safety profile of asapiprant given by PO or NG administration in patients ⩾ 50 years of age hospitalized with COVID-19 infection.

Other Objectives

- To investigate the effect of asapiprant on inflammation markers of COVID-19 infection, including but not limited to interleukin (IL)-6, C-reactive protein (CRP), IL-10, tumor necrosis factor alpha (TNF-α), interferon gamma (IFN-γ), interferon alpha (IFN-α), IFN-γ-induced protein 10 (IP-10), monocyte chemoattractant protein-1 (MCP-1), cluster of differentiation (CD)4+ and CD8+ T cells, and absolute lymphocyte count.
- Determine peak and trough concentrations of study drug after first dose and at steady state and, if possible, develop a population PK model.

2.2 Sample Size Determination

The sample size estimate of 312 total subjects is based on 31% progressing to RF by Day 28 in the placebo group and 15% progressing to RF in the asapiprant group using data from recently published COVID-19 studies (Guimaraes et al 2021, Temesgen et al, 2021). Using a Fishers Exact test with two-sided 5% Type 1 error and 90% power, the sample size to detect a 16% absolute improvement in the primary endpoint is 312 patients (1:1 randomization). Randomization will be stratified based on the patient's enrollment region (North America vs. South America) and age (50 to < 75 vs. ⩾ 75 years of age).

2.3 Study Endpoints and Evaluations

Primary Endpoint

- Proportion of patients who have died or progressed to RF as defined by progressing to the need for high-flow nasal cannula O_2 delivery, noninvasive positive pressure ventilation, invasive mechanical ventilation, or extracorporeal membrane oxygenation (ECMO) therapy by Day 28.

Secondary Endpoints

1. Safety as assessed by AEs as measured by the Common Terminology Criteria for Adverse Events (CTCAE) v 5.0
2. Proportion of patients surviving [Time Frame: at Day 14, at Day 28, at Day 57]
3. Proportion of subjects who survive without progression to RF [Time Frame: at Day 14, at Day 28, at Day 57]

4. Time to two successive negative viral titers in nasopharyngeal swab
5. Time to clinical worsening from baseline value (defined by time to ⩾ 1-point worsening on WHO Ordinal Scale for COVID-19):

World Health Organization Ordinal Scale for COVID-19

Patient State	Description	Score
Uninfected	No clinical or virological evidence of infection	0
Ambulatory	No limitation of activities	1
	Limitation of activities	2
Hospitalized, Mild Disease	Hospitalized, no oxygen therapy	3
	Oxygen by mask or nasal prongs	4
Hospitalized, Severe Disease	Noninvasive ventilation or high-flow oxygen	5
	Intubation and mechanical ventilation	6
	Ventilation + additional organ support – pressors, RRT, ECMO	7
Dead	Death	8

Abbreviations: ECMO = extracorporeal membrane oxygenation; RRT = renal replacement therapy

6. Proportion of patients who develop critical COVID-19 illness as defined by at least one the following:
    A. Respiratory failure defined based on resource utilization requiring at least one of the following: Endotracheal intubation and mechanical ventilation, oxygen delivered by high-flow nasal cannula (heated, humidified, oxygen delivered via reinforced nasal cannula at flow rates > 20 L/min with fraction of delivered oxygen ≥ 0.5), noninvasive positive pressure ventilation, ECMO clinical diagnosis respiratory failure (i.e., clinical need for one of the preceding therapies, but preceding therapies not able to be administered in setting of resource limitation)
    B. Hemodynamic compromise (defined by systolic blood pressure < 90 mm Hg, or diastolic blood pressure < 60 mm Hg or requiring vasopressors)
    C. Multi-organ dysfunction/failure
7. Time to clinical improvement from baseline value (defined by time to ≥ 1-point improvement on WHO Ordinal Scale for COVID-19 score – must be maintained through Day 28)
8. Mean change from baseline in WHO Ordinal Scale for COVID-19 score at Day 14/End of Treatment, Day 28, Day 57
9. Incidence and duration of intubation
10. Length of intensive care unit stay (ICU)
11. Incidence and duration of supplemental O_2 administration
12. Incidence and duration of noninvasive ventilation or high-flow nasal cannula O_2 administration

13. Incidence and duration of mechanical ventilation
14. Incidence and duration of mechanical ventilation plus additional organ support using vasopressors, and/or renal replacement therapy and/or ECMO
15. Daily ratio of oxygen saturation (SpO_2) to fractional inspired O_2 (SpO_2/FiO_2)
16. Length of hospital stay
17. Incidence of re-hospitalization through Day 57
18. Proportion of patients requiring intensive care unit admission post randomization

Exploratory Endpoints

1. Inflammation markers including: IL-6, CRP, IL-10, TNF-α, IFN-γ, IFN-α, IP-10, MCP-1, and absolute lymphocyte count
2. Assess peak and trough concentrations of asapiprant at steady state
3. Assess PGD2 pre-diagnostic to assess correlation with response to treatment for COVID-19 based on change in the WHO Ordinal Scale for COVID-19 score

Safety endpoints

1. Incidence and severity of treatment-emergent AEs (TEAEs)
2. Safety outcome measure:
    a. Occurrence of AEs
    b. Laboratory safety parameters such as complete blood count, blood glucose, electrolyte, hepatic and renal functions taken before the first dose and on Days 2, 5, 14 and 28 and the time of discharge to find any changes or any systemic effect after treatment

3 Analysis Sets

Enrolled Analysis Set: all patients who provided informed consent.

Intent-to-Treat (ITT): all patients who have been randomized. The ITT will be used for efficacy endpoints. Patients will be analyzed according to the treatment group to which they were randomized. The ITT will be used for the primary efficacy analysis.

Modified Intent-to-Treat population (mITT): all ITT patients with exclusion of randomized untreated patients and randomized ineligible patients.

Per-protocol Set (PPS): all mITT patients with further exclusion of patients with major protocol deviations related to data integrity or eligibility as judged by clinical/biostatistical study personnel prior to database lock. This population will be formed only if > 5% of patients would be excluded.

Safety Analysis Set (SAF): all patients who have received at least 1 dose of either asapiprant or placebo. The SAF will be used for safety analysis. Patients will be analyzed under actual treatment being received.

Pharmacokinetic (PK) Set: all patients who have received at least 1 dose of asapiprant and have at least 1 evaluable post-dose PK concentration value.

Pharmacodynamic (PD) Set: all patients who have received at least 1 dose of asapiprant and have at least 1 evaluable post-dose PD value.

The assignment of patients to the different analysis set will be done prior to unblinding during a blind data review meeting. All patients excluded from the different analysis sets will be listed together with the reason for exclusion.

The table below displays how different analyses will be applied to the different analysis sets:

Analysis	Enrolled	mITT	ITT	PPS	SAF	PK	PD
Disposition	X						
Protocol Deviation			X				
Exposure and compliance					X		
Demography and baseline characteristics			X				
Medication History			X				
Primary efficacy		X	X	X			
Secondary efficacy		X	X	X			
Safety					X		
PK						X	
PD							X

4 General Consideration for Data Analysis

All analyses will be conducted using SAS® version 9.3 or later (SAS Institute, Cary, NC 27513).

4.1 General Statistical Methods

The statistical analyses will be presented by treatment group for the different analysis sets as defined in section 3.

Summary tables will be structured with a column for each treatment in the order (asapiprant, Placebo, Total).

Continuous variables will be summarised using descriptive statistics, i.e., displaying number of patients in the respective analysis population, number of patients with data, mean, median, standard deviation (SD), minimum, and maximum.

Categorical variables will be summarised by using frequency counts and percentages. In addition, the number of patients with missing values will be displayed. Unless specified, the denominator for percentages will be the number of patients in the analysis set and treatment group under consideration.

Means and medians will be reported at 1 more significant digit than the precision of the data. Standard deviations and confidence intervals (CIs) will be reported at 2 more significant digits than the precision of the data. Minimum and maximum will be reported to the same level of precision as the original observations. Percentages will be presented to 1 decimal place if not otherwise stated.

If the number of patients in a category is 0, then percentage will not be displayed, and only a count of 0 will be shown.

All p-values will be two-sided and reported to 4 decimal places at least. Values less than 0.0001 will be displayed as <0.0001.

In listings, data will be sorted by treatment, site, and patient ID, and when appropriate by visit or other identifiers for sequence or type of observation.

The confidence level for calculation of CIs will be chosen as (1-significance level) of the respective statistical test.

If a qualitative parameter in a table has missing values, a row with the number of missing values will be added.

Factor 365.2425 is used for duration conversions between days and years. Factor 30.436875 is used for duration conversions between days and months.

4.2 Covariates and Strata

The primary endpoint (proportion who have died or progressed to RF at 28) will be analyzed using a logistic regression with treatment group and covariates, age group (50 to < 75 years, and ≥75 years), and region. Other covariates including previous COVID-19 therapy may be explored as exploratory analyses.

4.3 Subgroups

The following sub-groups will be examined for the primary endpoint. Results will be displayed in both tabular form. Subgroup analyses are exploratory.

- Age (50 to < 75 vs ≥ 75)
- Gender (male vs female)
- previous COVID-19 therapy (yes vs no)
- Regions (North America vs South America)
- Baseline WHO ordinal clinical scale (3 or 4)

4.4 Missing Data

In general, missing data will not be imputed except partial and missing dates as specified below. The data will be analyzed as they are recorded.

Details of handling missing primary endpoint are addressed in section 10.1 primary analysis.

WHO Ordinal Scale for COVID-19 scores will be analyzed by carrying forward any early deaths to Day 28.

Partial or missing dates and times

The missing component(s) of incomplete dates (e.g., start and/or stop dates of AE, concomitant medication, medical history) will be assumed as the most conservative value possible. For example, if the start date has a missing day value, the first day of the month will be imputed for study day computations, etc. If day is missing for an end date, the last day of the month will be imputed. If the start date has a missing month value, the first month of the year will be imputed for study day

computations, etc. If month is missing for an end date, the last month of the year will be imputed. For determination of treatment-emergent status, the start date will be imputed as the date of the first dose of study drug, unless there is clear evidence (through comparison of partial dates/times) to suggest otherwise.

Date imputation will only be used for computational purposes such as treatment-emergent status, etc. Actual date values, as they appear in the original electronic case report forms (eCRFs), will be presented in the patient data listings.

4.5 Visit Windows

The following table outlines the analysis visits that data will be mapped to for analysis of efficacy, safety and pharmacodynamics data collected during the study. If after mapping, multiple data points share the same analysis visit, the data point closest to the expected study visit will be used. If there is a tie, the earlier of the visits will be used.

Study Visit	Analysis Visit	Analysis Window
Screening	Screening	-3 to -1
Days 1 – 13	Days 1 – 13	0
Days 14	Days 14	+2 day
Day 28 Follow-up	Day 28	± 2 day
Day 57 Phone Follow-up Visit	Day 57	± 4 day

4.6 Study Days

Study day is defined as the number of days since first dose of IMP and, for a particular date, and is calculated as:

> Study day = Assessment date – Date of first dose of IMP + 1, if date of assessment is on or after date of first dose of IMP
>
> or
>
> Study day = Assessment date – Date of first dose of IMP, if date of assessment is prior to the date of first dose of IMP

Therefore, the date of the first dose of IMP will be Day 1.

In the situation where the assessment date is partial or missing, Study Day, and any corresponding durations will appear partial or missing in the listings.

4.7 Definition of Baseline Values

The baseline value is defined to be the last non-missing assessment prior to randomization. For baseline WHO ordinal scale, if not collected prior to randomization, the earliest collected WHO ordinal scale collected on the same day as randomization will be used as baseline. Pre-randomization assessments on the same day as randomization will be considered the baseline assessment for each of these respective study procedures. Unless otherwise specified, change from baseline calculations for a treatment window assessment will be the applicable treatment window assessment minus the

baseline assessment. If either the treatment window assessment value or the baseline value is/are missing, then change from baseline will be set to missing.

Age in years will be calculated as (date of informed consent – January 01 of year of birth)/ 365.2425.

4.8 Planned Analyses

The following statistical analyses are planned for this study:

- Unblinded safety analyses for Data Monitoring Committee (DMC): An independent DMC will be utilized throughout the study to monitor patient safety. DMC meetings for safety will be conducted after 33, 100, and approximately 208 subjects complete assessments through the 28-day follow up (Day 28), approximately 14 days after the final dose of study drug. A DMC charter will be developed preceding study initiation and will define DMC processes and meetings, and required report contents.
- Interim analysis: An interim look for safety and primary endpoint will be conducted by the DMC when approximately 208 patients in the ITT population complete assessments through the 28-day follow up (Day 28), approximately 14 days after the final dose of study drug.

  When all subjects enrolled under the targeted or adjusted sample size have completed the study through Day 28 (i.e., full study enrolment is completed), the clinical database will be cleaned, frozen, and an assessment of the primary efficacy endpoint and secondary, exploratory, and safety endpoints up to Day 28 in the ITT population will be created. Individuals involved with the conduct of the study will not have access to unblinded data until database lock. Final analysis: final analysis on efficacy and safety will occur after the database is locked.

4.9 Multiple Comparison/Multiplicity

No multiplicity adjustment is planned.

5 Study Population and Disposition

5.1 Analysis Population

The number and percentage of patients in each analysis set as defined in section 3 will be presented overall and by randomised treatment. The denominators for percentages will be based on the number of enrolled patients by randomized treatment and overall.

Overall patient status will be provided in a data listing, containing analysis population assignment, randomized treatment, actual treatment, date of consent, date of first dose, date of termination.

5.2 Patient Disposition

The number and percentage will be presented by randomized treatment and overall for:

- patients screened, randomized
- patients completing the study drug treatment
- patients completing the study
- patients who prematurely discontinued study drug treatment
- patient who prematurely withdrew from the study

- reason for discontinuation of study drug treatment
- reason for withdrawing from the study as recorded in eCRF

The denominators for percentages will be based on the number of patients in ITT by randomized treatment and overall, except for patients screened and patients randomized.

Discontinued patients will be provided in a data listing by patient.

5.3 Protocol Deviations

Major protocol deviations will be judged by clinical/biostatistical study personnel prior to database lock.

Major protocol deviations may include but not limited to followings:

- Deviations of inclusion and exclusion criteria
- Receiving the wrong study intervention
- Non-compliance with study intervention defined as the participant taking < 80% of the study intervention through their last day of dosing.
- Taking a prohibited medication
- Other

The number and percentage of patients with major deviations will be summarized. The number of deviations by type of deviation will be summarized by treatment and overall. The denominators for percentages will be based on the number of patients in ITT population by randomized treatment and overall.

All protocol deviations (both major and minor) will be provided in a data listing by patient.

5.4 Inclusion/Exclusion Criteria

Inclusion and exclusion criteria deviations will be provided in a data listing by patient.

6 Demographics and Disease Characteristics

6.1 Demographics and other baseline characteristics

The following demographic characteristics will be summarized by treatment group and overall for patients in ITT population using descriptive statistics. No statistical hypothesis tests will be performed.

- Sex (male, female)
- Age (years)
- Age group (50 to < 75, ≥ 75 years)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Unknown, Other)
- Ethnicity (Hispanic or Latino, not Hispanic or Latino, not reported, unknown)
- Weight (kg) at baseline
- Height (cm) at baseline

- BMI (kg/m^2) at baseline
- SARS-CoV-2 (COVID-19) history
- PGD2 result at baseline
- Virus serology result of presence of HIV, Hepatitis B surface antigen (HBsAg), and HCV antibody if collected in eCRF.

Demographics and other baseline characteristics will be provided in a data listing by patient.

6.2 Medical History

The medical conditions are coded according to Medical Dictionary for Regulatory Activities (MedDRA® version 23.1) and will be classified as follows:

- Previous medical conditions, if medication is not "Ongoing at study entry" as indicated in CRF
- Concomitant medical conditions, if medication is "Ongoing at study entry" as indicated in CRF

The frequency and percentage of medical conditions recorded from medical history will be presented after classification into previous and concomitant conditions by system organ class (SOCs) and preferred terms (PT) by randomized treatment and overall. If patients have more than one disease within an SOC or PT, they will be counted only once for the respective SOC or PT.

The denominators for percentages will be based on the number of SAF by treatment group and overall.

7 Exposure and Compliance

7.1 Extent of Exposure

Duration of study drug exposure, defined as the number of days between the date of first dose following randomization and the date of the last dose, will be summarized by treatment group and over for the patients in SAF.

Duration will be calculated as date of last dose – date of first dose + 1.

If the date of first intake of IMP is missing, it will be assumed to be identical to the date of randomization.

7.2 Compliance

Compliance will be calculated and summarized by treatment group and overall. Compliance will be calculated as 100 × (total number of tablets taken / total number of tablets expected to be taken). The number of tablets taken will be derived as the number of tablets dispensed minus the sum of the number of tablets returned and the number of tablets reported lost/missing based on the CRF reported doses taken in hospital and the dose log. The percent compliance will be summarized. Summaries will be presented for SAF.

A listing of drug accountability will be presented for all medication distributed to each patient, including number of tablets dispensed, returned, lost/stolen, taken, and overall percent compliance.

8 Prior and Concomitant Medication

Prior and concomitant medications are coded according to the World Health Organization drug dictionary (WHO-Drug version MAR, 01, 2019) and stored with ATC codes and generic names.

Medication will be classified as prior if the stop date and/or time was before the first dose of IMP. All other medications are defined as concomitant. Missing or partly missing stop dates will be imputed using the rules defined in Section 4.4.

The number and frequency of previous and concomitant medications will be given per ATC level 2 and preferred term by treatment and overall. If a patient has received more than 1 drug within an ATC class/preferred term, he/she will be counted only once for the ATC class/preferred term.

9 Prior and Concomitant Procedure

Procedure will be classified as previous if the stop date and/or time was before the first dose of IMP. All other procedures are defined as concomitant. Type of procedure (prior or concomitant) and indication of procedure will be summarized. All summaries will be presented by dose cohort and overall.

Prior and concomitant procedure information will also be provided in a data listing.

10 Efficacy

10.1 Primary efficacy analysis

The number and percentage of patients who have died or have progressed to RF through Day 28 will be presented by treatment. Hypothesis test comparison of proportion between treatment will be analyzed using a logistic regression with predefined covariates and treatment group. The predefined covariates are region and age group. Other covariates including previous COVID-19 therapy may be explored as exploratory analyses. Estimation of the covariate-adjusted risk difference will be reported using Cochran-Mantel-Haenszel (CMH) weighting and estimates from the logistic regression as described by Ge and colleagues (Ge et al., 2011, Appendix C). Two-sided p-value of global test comparison between active and placebo will be reported. Odds ratio of each covariate and two-sided 95% CI will also be reported. This analysis will be repeated without the covariates.

The denominator for the proportions alive and RF-free at Day 28 will include all randomized subjects.

Time from date of randomization to death or first RF through Day 28, whichever comes first, will also be displayed using the Kaplan-Meier estimator and analyzed using a proportional hazards model analysis with the same covariates. If a patient doesn't have documented death or RF, it will be censored at the Day 28. For subject who discontinued study before Day 28 without documented outcome, it will be censored at the date of discontinuation. This analysis will be repeated without the covariates. Test statistics will be the Hazard Ratio with 95% CIs and p-values. Sensitivity test will be done using a Fishers Exact Test.

Primary endpoint will be analyzed using ITT, mITT, PPS (if formed) analysis set under randomized treatment.

Primary endpoint analysis will be repeated for subgroups as specified section 4.3.

Sensitivity analysis and handling of missing primary endpoint data

If $\leqslant$ 5% of subjects lack sufficient data to confirm the primary endpoint, four sensitivity analyses will be performed on logistic regression: 1) treat any missing primary outcomes as successes; 2) replaces

any missing primary outcomes as failures; 3) missing primary outcome imputed as success for drug and failure for controls; 4) missing primary outcome imputed as failure for drug and success for controls.

If more than 5% of subjects have insufficient data to confirm RF-free and alive at Day 28, BioAge will employ a MI strategy (SAS MI, MIANALYZE) with missing at random (MAR) based protocol stratification factors (age and region), treatment, and the discharge status prior to or on Day 28. Following steps will be followed for multiple imputation.

Step 1: which will be carried out to produce n (exp. 25) imputed datasets

The SAS procedure PROC MI using the FCS (fully conditional specification method) statement with LOGISTIC imputation method to impute RF28 assuming no covariate will be missing.

One must also specify a VAR statement with outcome variable being the last in the order. If any covariate is expected to have missing and need to be imputed as well "LOGISTIC(RF28SUB_MI=TRT01P REGION AGEGRN DISC28" also need to be added after the FCS statement.

```
PROC MI DATA=RF28_DATA SEED=1305417 NIMPUTE=25 OUT=OUTLOG;
      CLASS RF28 TRT01P REGION AGEGRN DISC28;
      VAR TRT01P REGION AGEGRN DISC28 RF28;
      FCS LOGISTIC(RF28SUB_MI=TRT01P REGION AGEGRN DISC28) ;
RUN;
```

Where:

RF28 is respiratory failure or death prior to or on Day 28

TRT01P is randomized treatment (Placebo/Asapiprant)

REGION: region of patient's location (North America/South America)

AGEGRPN: age group (50 - <75, >=75)

DISC28: discharge status prior to or on Day 28

Step 2: Each of 25 complete data set will be analyzed using standard logistic analyses as specified for primary efficacy analysis in SAP, i.e. PROC LOGISTIC will be used for RF28 with predefined covariates (region and age group) and treatment group.

```
PROC LOGISTIC DATA=OUTLOG ;
   CLASS RF28 TRT01P REGION AGEGRN ;
   MODEL RF28=TRT01P REGION AGEGRN / COVB;
   BY  IMPUTATION ;
   ODS OUTPUT PARAMETERESTIMATES=LGSPARMS;
RUN;
```

Step 3: The results logistic regression from the 25 complete data sets are combined for the parameter inference (i.e. coefficient estimate, 95% confidence interval, and p-value) using PROC MIANALYZE

```
PROC MIANALYZE PARMS=LGSPARMS;
   MODELEFFECTS INTERCEPT TRT01P REGION  AGEGRN;
RUN;
```

10.2 Secondary efficacy analyses

All secondary efficacy endpoints will be analyzed without adjustment for multiplicity using ITT and mITT and PPS (if formed) under randomized treatment based on the data observed. No missing data will be imputed unless specified otherwise.

10.2.1 Proportion of binary endpoints

The number and percentage of the following binary endpoints will be presented by treatment and timepoint. Comparison of proportion of patients between treatment at each timepoint will be analysed by the same manner as primary endpoint.

- *Proportion of patients who had survived at each timepoint at Day 14, at Day 28, at Day 57:* which is defined as patients who had no documented death prior or on timepoint at Day 14, 28, 57.
- *Proportion of patients who survive without progression to RF at each timepoint at Day 14, Day 28, Day 57*: which is defined as patients who had neither documented death nor respiratory failure prior or on timepoint at Day 14, 28, 57.
- *Proportion of patients who had intubation during the study:* which will be defined as proportion of patients who had any documented intubation during the study.
- *Proportion of patients who develop critical COVID-19 illness as defined by at least one the following*:
    A. Respiratory failure defined based on resource utilization requiring at least one of the following: Endotracheal intubation and mechanical ventilation, oxygen delivered by high-flow nasal cannula (heated, humidified, oxygen delivered via reinforced nasal cannula at flow rates > 20 L/min with fraction of delivered oxygen $\geqslant$ 0.5), noninvasive positive pressure ventilation, ECMO clinical diagnosis respiratory failure (i.e., clinical need for one of the preceding therapies, but preceding therapies not able to be administered in setting of resource limitation)
    B. Hemodynamic compromise (defined by systolic blood pressure < 90 mm Hg, or diastolic blood pressure < 60 mm Hg or requiring vasopressors)
    C. Multi-organ dysfunction/failure
- *Proportion of patients requiring intensive care unit admission*: which is defined as patients who had been admitted to intensive care unit during the study.
- *Incidence of supplemental O_2 administration*: which will be defined as proportion of patients who had any documented post-dosing supplemental O_2 administration during the study.
- *Incidence of noninvasive ventilation or high-flow nasal cannula O_2 administration*: which will be defined as proportion of patients who had any documented post-dosing noninvasive ventilation or high-flow nasal cannula O_2 administration.
- *Incidence of mechanical ventilation*: which will be defined as proportion of patients who had any documented post-dosing mechanical ventilation.
- *Incidence of mechanical ventilation plus additional organ support using vasopressors, and/or renal replacement therapy and/or ECMO*: which will be defined as proportion of patients who had any documented post-dosing mechanical ventilation plus additional organ support using vasopressors, and/or renal replacement therapy and/or ECMO.

- *Incidence of re-hospitalization through Day 57*: which will be defined as proportion of patients who are hospitalized again after the discharge of first hospitalization.

10.2.2 Time to event endpoints

Following time to event endpoints will be presented by treatment group, using the Kaplan-Meier estimator. A Cox proportional hazards model with treatment, previous COVID-19 therapy (yes/no), region, and age group (50 to < 75 vs. ≥ 75 years of age) as covariates will be used to estimate the hazard ratio and compare treatment groups. In addition, the log rank test and stratified log rank test, with age group and region as the strata, will be provided.

- *Time to two successive negative viral titers in nasopharyngeal swab*: which is defined as time from the date of randomization to the date of second post-dosing viral titers collection with negative result of two successive negative viral titers tests. If a patient doesn't have documented two successive negative viral titers, it will be censored at the date of Day 28 or date of discontinuation from study, which ever comes earlier.
- *Time to clinical worsening from baseline value (defined by time to ⩾ 1-point worsening on WHO Ordinal Scale for COVID-19):* which is defined as time from the date of randomization to the date of first post-dosing collection of WHO Ordinal Scale for COVID-19 showing higher than baseline, i.e., ⩾ 1-point worsening. If a patient doesn't have documented worsening post-baseline WHO Ordinal Scale, it will be censored at the date of Day 57 or date of discontinuation from study, which ever comes earlier.
- *Time to clinical improvement from baseline value (defined by time to ≥ 1-point improvement on WHO Ordinal Scale for COVID-19 score – must be maintained through Day 28):* which is defined as time from the date of randomization to the date of first post-dosing collection of WHO Ordinal Scale for COVID-19 showing a decrease from baseline, i.e., ≥ 1-point decrease demonstrating clinical improvement. If a patient doesn't have documented improving post-baseline WHO Ordinal Scale, it will be censored at the date of Day 28 (including death prior to Day 28) or date of discontinuation from study, whichever comes earlier.

10.2.3 Mean change in WHO Ordinal Scale for COVID-19 score at Day 14/End of Treatment, Day 28, Day 57

Mean change in WHO Ordinal Scale at Day 14, 28, and 57 will be analyzed using ANCOVA on change at each time point. The model will include the stratification factors. The least squares means (LS-means) for each treatment, LS-means treatment difference and corresponding CI for the LS-means difference and p-value for the treatment difference, will be reported at the Day14, 28, and 57.

Results from Wilcoxon rank sum test and t-tests will also be reported for change in WHO Ordinal Scale at the Day 14, 28, and 57.

10.2.4 Other continuous secondary efficacy endpoints

Following continuous endpoints will be analyzed using an analysis of covariance (ANCOVA) model, with treatment as a fixed effect and baseline WHO Ordinal Scale score, age group (50 to < 75 vs. ≥ 75 years of age) and region. Results from Wilcoxon rank sum test and t-tests will also be reported. The ITT and PP analysis (if formed) sets will be used for these exploratory subset analyses.

- *Duration of intubation:* which will be defined as duration of first post-dosing intubation based on start and end date of the intubation.

- *Length of intensive care unit (ICU) stay in days* which will be defined as duration of first post-dosing ICU stay based on start and end date/time of the stay.
- *Duration of supplemental O_2 administration in days* : which will be defined as duration of first post-dosing supplemental O_2 administration based on start and end date/ of the O_2 administration.
- *Duration of noninvasive ventilation or high-flow oxygen therapy in days*: which will be defined as duration of first post-dosing noninvasive ventilation or high-flow oxygen therapy based on start and end date/time of the noninvasive ventilation or high-flow oxygen therapy.
- *Duration of mechanical ventilation in days*: which will be defined as duration of first post-dosing mechanical ventilation based on start and end date/time of the mechanical ventilation.
- *Duration of mechanical ventilation plus additional organ support using vasopressors, and/or renal replacement therapy and/or ECMO*: which will be defined as duration of first post-dosing mechanical ventilation plus additional organ support using vasopressors, and/or renal replacement therapy and/or ECMO based on start and end date/time of the mechanical ventilation plus additional organ support using vasopressors, and/or renal replacement therapy and/or ECMO.
- *Daily ratio of oxygen saturation (SpO_2) to fractional inspired O_2 (SpO_2/FiO_2):* which will be defined as SpO_2/FiO_2 as collected in CRF. The ratio will be analyzed for each day with the value being collected. If FiO_2 was not collected for some O_2 delivery methods, following guide can be followed to convert O_2 flow rate to FiO_2.

O_2 delivery method	O_2 flow rate (l/min)	Estimated FiO_2
Nasal cannula ("NASAL CANNULA LARGE TUBING" and "NASAL CANNULA TRADITIONAL TUBING"	1	24
	2	28
	3	32
	4	36
	5	40
	>=6	44
Nasopharyngeal catheter	1	25
	2	30
	3	35
	4	40
	5	50
	>=6	60
Face mask (simple mask)	<5	Follow nasal cannula
	5	40
	6-<7	50
	>=7	60

Face mask with reservoir (non-rebreather face mask)	<6	Follow nasal cannula
	6	60
	7	70
	8	80
	9	90
	10	95
	>10	100
Venturi mask	1	22
	2	24
	3	26
	4	28
	6	31
	8	40
	10	50
	>10	60

In case of fractional value of flow rate (0.5, 1.5, 2.5, 3.5, 4.5), FiO_2 = (Oxygen Flow Rate *4)+20 to avoid rounding.

- *Length of hospital stay, after start of study treatment, in days*: which will be defined as duration from date of randomization to the date of first discharge.

11 Safety

All safety endpoints will be analyzed using the SAF under received treatment based on the data observed. No missing data will be imputed.

11.1 Adverse events

Adverse events will be coded using MedDRA version 23.1 and will be graded according to the NCI-CTCAE version 5.0. AE summaries will be presented by primary SOC and PT.

Treatment-emergent Adverse events will be considered as any event or worsening of event with onset on or after receiving the first dose of study drug through Day 57.

Summaries will be presented by actual treatment and overall using the SAF. Although a PT or SOC may be reported more than once for a patient, each patient will only be counted once.

Number and frequencies of patients with TEAEs as well as number of events will be given by SOC and by PT within each SOC for the following:

- All TEAEs by treatment group
- Serious TEAEs by treatment group
- All TEAEs considered related by treatment group
- Serious TEAEs considered related by treatment group

- All TEAEs by maximum severity and by treatment group
- TEAEs leading to death by treatment group
- TEAEs leading to study discontinuation of study drug by treatment group

Additionally, number and frequencies of patients with TEAEs as well as number of events will be presented by decreasing frequency for all TEAEs, serious TEAEs and TEAEs considered at least possibly related.

If severity is missing, the event will be graded as severe unless the event being recorded is leading to life-threatening or death and will be graded accordingly as grade 4 (life-threatening) or 5 (death). If relationship to study drug is missing, the event will be assessed as unrelated if it started before start of IMP; in all other cases it will be assumed to be related.

Patient listings will be provided for patients with adverse events, treatment emergent adverse events, serious adverse events, for all patients with adverse events leading to withdrawal of IMP, and for all adverse events leading to death.

11.2 Deaths

Adverse events and the principal cause of death will be coded using MedDRA version 23.1.

Deaths will be summarized by MedDRA primary system organ class and preferred term of the principal cause of death.

11.3 Vital signs

Vital signs measurements (including blood pressure, temperature, respiratory rate, pulse rate, SpO_2, and ratio of SpO_2/FiO_2) and change from baseline values will be summarized over time with n, mean, SD, and median, minimum, and maximum. Summaries will be presented by treatment and overall.

Vital signs information will also be provided in a data listing.

11.4 Clinical safety laboratory

Haematology, coagulation, and biochemistry

The following statistical analyses will be presented:

- Quantitative data will be examined using descriptive statistics (number of patients with data, mean, SD, median, minimum, and maximum) of actual values and changes from baseline to each visit over time.
- Qualitative data based on reference ranges will be described according to the categories (i.e., Low, Normal, High).
- Shift tables showing changes with respect to the normal range between baseline and each timepoint as well as the worst value at any post-baseline visit.

A separate listing will also be provided. Laboratory values that are outside the normal range will also be flagged in the data listings, along with corresponding normal ranges.

Urinalysis

The following statistical analyses will be presented only at screening:

- Quantitative data will be summarized using descriptive statistics (number of patients with data, mean, SD, median, minimum, and maximum) of actual values.
- Qualitative data based on reference ranges will be described according to the categories (i.e., Low, Normal, High).

11.5 ECG

Baseline and changes from baseline in RR, PR, QT, QTc, QTcF, QRS and HR will be summarized at each timepoint using descriptive statistics Descriptive statistics by treatment and timepoints (mean, SD, 25th percentile, median, 75th percentile, minimum, and maximum)

Number and percentage of overall interpretation of electrocardiogram (ECG) (i.e., normal, abnormal not clinically significant, abnormal clinically significant) and shift from baseline will be summarized by treatment and overall.

11.6 Physical examination

A complete physical examination (PE) will be performed at Screening. After Screening, abbreviated physical examination will be performed at Days 5 and 14, at discharge, and at Day 28.

Number and percentage of result (normal, abnormal, not done) of full PE result at screening and post-baseline abbreviated PE result at each visit will be summarized by treatment and overall.

PE will be provided in a data listing by patient.

12 Other Analyses – please specify

The pharmacokinetic (PK) analyses will be described in a separate analysis plan.

For inflammation markers of COVID-19 infection, including but not limited to IL-6, CRP, IL-10, TNF-α, IFN-γ, IFN-α, IFN-γ, IP-10, MCP-1, and absolute lymphocyte count, summaries will be presented by treatment and overall using descriptive statistics. Both the observed value and the change from baseline value will be summarized.

13 Interim Analyses

An interim analysis for safety and primary endpoint will be conducted after the first approximately 208 patients complete assessments through the 28-day follow up (Day 28), approximately 14 days after the final dose of study drug. Unblinded interim analysis results will be reviewed by the independent and unblinded DMC for safety evaluation. Details of the roles/responsibilities of DMC members and each party, meeting schedule and intended activities are described in a separate DMC Charter.

Analysis of primary endpoint at the interim will follow as described in section 10.1.

An α spend is not incorporated into the sample size to account for the interim, as there is no potential for stopping for superiority.

Efficacy analyses at interim analysis (IA) may yield Conditional Power (CP) results in the Mehta and Pocock promising zone (0.34, 0.90) (Mehta and Pocock, 2011). If so, to maintain the statistical power of the study and to avoid any type 1 error penalty, the Sponsor will consider the option of increasing the sample size (according to the rules provided by Mehta and Pocock), not to exceed N=624 (twice the planned sample size). For this planned IA, Z-score evidence based on logistic regression as 10.1

at the IA of between Z=1.41 and 2.21 would define the promising zone and opportunity for BioAge to expand the study. For example, if at IA, Z=1.8 (CP=0.66), then to maintain the power at 0.90 the study could be expanded from 312 subjects to 500. For results in the Mehta and Pocock unfavorable zone, or in the favorable zone, the study will continue to conclusion at N=312 subjects.

14 Statistical Analyses for Safety Monitoring

The statistical analysis for safety monitoring is specified in a separate DMC charter.

15 References

1. Guideline on adjustment for baseline covariates in clinical trials, EMA/CHMP/295050/2013, 2015
2. Guimaraes PO, Quirk D, Furtado RH, et al. Tofacitinib in Patients Hospitalized with Covid-19 Pneumonia. N Engl J Med 2021;385(5):406-415. DOI: 10.1056/NEJMoa2101643.
3. Senn SJ. Covariate imbalance and random allocation in clinical trials. Stat Med 1989;8:467e75.
4. ICH Harmonised Tripartite Guideline E9, Statistical Principles for Clinical Trials, 1998
5. ICH Harmonised Tripartite Guideline E3, Structure and Content of Clinical Study Reports. 1996
6. ICH Harmonised Tripartite Guideline E6(R2), Guideline for Good Clinical Practice, 2018
7. ICH guideline E2F, Note for guidance on development safety update reports, EMA/CHMP/ICH/309348/2008, September 2010
8. Guideline on missing data in confirmatory clinical trials, EMA/CPMP/EWP/1776/99 Rev. 1, 2010
9. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical 6 principles for clinical trials, Step 2b (Draft), EMA/CHMP/ICH/436221/2017, 2017
10. Mehta, C.R., and Pocock, S.J. (2011) Adaptive increase in sample size when interim results are promising: a practical guide with examples. Stat Med 40, 3267–3284.
11. Temesgen Z, Burger CD, Baker J, et al. Lenzilumab Efficacy and Safety in Newly Hospitalized Covid-19 Subjects: Results from the Live-Air Phase 3 Randomized Double-Blind Placebo-Controlled Trial. medRxiv 2021. DOI: 10.1101/2021.05.01.21256470.

16 Addendum

Based on the following information and justifications, baseline WHO ordinal scale is removed from primary endpoint logistic regression:

1) enrollment stopped before interim analysis and all enrolled patients have passed Day 28;

2) based blinded data review, all those subjects who had baseline WHO scale=3 had no respiratory failure, and therefore leading quasi-complete separation on primary endpoint.