Statistical Analysis Plan

Study ID: 214367

Official Title of Study: A randomized, multi-center, double-blind, placebo-controlled study to assess the safety and efficacy of monoclonal antibody VIR-7831 for the early treatment of coronavirus disease 2019 (COVID-19) in outpatients.

NCT ID: NCT04545060

Date of Document: 29-JAN-2021

Division	:	Worldwide Development
Information Type	:	Statistical Analysis Plan (SAP)

Title Page

Protocol Title: A randomized, multi-center, double-blind, placebo-controlled study to assess the safety and efficacy of monoclonal antibody VIR-7831 for the early treatment of coronavirus disease 2019 (COVID-19) in outpatients

Protocol Number: VIR-7831-5001 (GSK Study 214367) Original

Compound Number: VIR-7831 (GSK4182136)

Short Title: VIR-7831 for the early treatment of COVID-19 in outpatients
Sponsor Name: This study is sponsored by Vir Biotechnology, Inc.

GlaxoSmithKline is supporting Vir Biotechnology, Inc. in the conduct of this study.

Vir Biotechnology, Inc.
499 Illinois Street, Suite 500
San Francisco, CA 94158
USA

Regulatory Agency Identifier Number(s)

Registry	**ID**
IND:	IND-149315
EudraCT:	2020-002871-36
NCT:	NCT04545060

SAP Author(s):

Author	Date
PPD Statistics Leader, Biostatistics, GlaxoSmithKline	29 JAN 2021

SAP Team Review Confirmations (E-mail Confirmation):

Approver	Date of E-mail Confirmation
PPD PPD, Clinical Research, Vir Biotechnology Inc.	29 JAN 2021
PPD PPD Pharma Safety, GlaxoSmithKline	29 JAN 2021
PPD Programmer/Analyst, Clinical Programing, GlaxoSmithKline	29 JAN 2021

Clinical Statistics Line Approval (E-Mail):

Approver	Date
PPD PPD, Biometrics, Vir Biotechnology Inc	29 JAN 2021
PPD PPD, Biostatistics, GlaxoSmithKline	29 JAN 2021

TABLE OF CONTENTS

PAGE

Version history

This Statistical Analysis Plan (SAP) for study VIR-7831-5001 (GSK Study 214367) is based on the original protocol dated 27-JUL-2020 and protocol amendment dated 23-DEC-2020.

Table 1 SAP Version History Summary

SAP Version	Approval Date	Change	Rationale
3.0	Current Version	The key changes include: • Amendment of the hierarchy strategy for the secondary endpoints for FluPRO Plus • Clarification of handling of missing data for FluPRO Plus	Following partial use of paper PROs, additional clarification around handling of missing data required.
2.0	14-Jan-2021	The key changes include: • Amendment of the alpha-spending function from O'Brien Flemming to Hwang-Shih-DeCani ($\gamma = 1$) rules for overwhelming efficacy following emergent data on neutralising mAbs • Due to increasing importance of gender as a prognostic factor for progression, switching from a stratified CMH test to a log-binomial model including age, duration of symptoms, region and gender as covariates • Amendment of the hierarchy strategy for the secondary endpoints following emergent data on neutralising mAbs • Inclusion of a patient reported symptoms sustained response based on the adapted Flu-PRO Plus • Other minor, grammatical and	Update to SAP in response to regulatory feedback, clarifications and emergent data in the early treatment of COVID-19 with neutralising mAbs

		typographical corrections to improve readability	
1.0		Not Applicable	Original version

1. INTRODUCTION

The purpose of this SAP is to describe the planned analyses to be included in the Clinical Study Report for Study VIR-7831-5001. Details of the planned interim analyses, in addition to the final Day 29 efficacy analyses and 24-Week Safety Follow-Up analyses, are provided.

Descriptive study population analyses, such as summary of demography and baseline characteristics, and additional detail with regards to data handling conventions and the specification of data displays will be provided in the Output and Programming Specification (OPS) document.

1.1. Changes to the Protocol Defined Statistical Analysis Plan

There were no changes or deviations to the planned statistical analysis specified in the protocol amendment 1 (Dated: 22-DEC-2020).

1.2. Objectives, Endpoints and Estimands

1.2.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
• Evaluate the efficacy of VIR-7831 versus placebo in preventing the progression of mild/moderate COVID-19	• Proportion of participants who have progression of COVID-19 through Day 29 as defined by: 1. Hospitalization > 24 hours for acute management of illness OR 2. Death
Secondary	
Safety • Determine the safety and tolerability of VIR-7831 compared to placebo	• Occurrence of adverse events (AEs) • Occurrence of serious adverse events (SAEs) • Occurrence of adverse events of special interest (AESI)
• Assess the immunogenicity of VIR-7831	• Incidence and titers (if applicable) of serum ADA to VIR-7831
Pharmacokinetics • Assess the pharmacokinetics (PK) of VIR-7831 in serum	• VIR-7831 pharmacokinetics (PK) in serum
Efficacy • Evaluate the efficacy of VIR-7831 versus placebo in preventing the progression of mild/moderate COVID-19	Proportion of participants who have progression of COVID-19 through Day 29 as defined by: 1. Visit to a hospital emergency room for management of illness OR 2. Hospitalization for acute management of illness OR Death
• Evaluate the impact of VIR-7831 versus placebo on the duration and the severity	• Mean change in FLU-PRO Plus total score comparing Vir-7831 vs. Placebo

Objectives	Endpoints
of COVID-19 clinical symptoms	(AUC through Day 7) • Time to symptom alleviation using the FLU-PRO Plus
• Evaluate the efficacy of VIR-7831 versus placebo in reducing SARS-CoV-2 viral load	• Change from baseline in Viral load in nasal secretions by qRT-PCR at Day 8
• Evaluate the efficacy of VIR-7831 versus placebo in preventing COVID-19 respiratory disease progression	• Proportion of participants who progress to develop severe and/or critical respiratory COVID-19 as manifest by requirement for and method of supplemental oxygen at Day 8, Day 15, Day 22, or Day 29
• Evaluate the efficacy of VIR-7831 versus placebo in preventing mortality	• 29-day, 60-day, and 90-day all-cause mortality
Exploratory	
• Evaluate the impact of VIR-7831 versus placebo on incidence and duration of time on a ventilator, ICU length of stay (LOS), and total hospital LOS	• Total number of ventilator days from randomization through 90 days • Total intensive care length of stay (LOS) • Total hospital LOS
• Evaluate the efficacy of VIR-7831 versus placebo in preventing hospitalization from non-respiratory complications of COVID-19	• Proportion of participants hospitalized due to non-respiratory complications of COVID-19 (cardiac, renal, neurologic, hematologic events) by Day 8, Day 15, Day 22, and Day 29
• Monitor on-treatment emergence of SARS-CoV-2 resistant mutants against VIR-7831	Emergence of viral resistance mutants to mAb by SARS-CoV-2
• Evaluate the efficacy of VIR-7831 versus placebo in reducing duration of SARS-CoV-2 viral shedding and viremia	• Detection of SARS-CoV-2 in nasal secretions by RT-PCR at baseline and during follow-up period through Day 29 • Detection of SARS-CoV-2 in blood by qRT-PCR RT-PCR at baseline and during follow-up period through Day 29
• Evaluate the efficacy of VIR-7831 versus placebo in reducing SARS-CoV-2 viral load	• Change from baseline in viral load in nasal secretions and blood by qRT-PCR
• Evaluate the incidence of respiratory viral co-infection with SARS-CoV-2	• Respiratory pathogen detection in nasal secretions by PCR on Day 1
• Evaluate the effect of VIR-7831 versus placebo on potential biomarkers of host response to SARS-CoV-2	• Evaluation of host immune responses and exploratory biomarkers related to SARS-CoV-2 and/or VIR-7831, including genetic, cellular, transcriptomic, and proteomic parameters
• Measure the impact of VIR-7831 treatment on health-related quality of life	• Change from baseline in Work Productivity and Activity Impairment

Objectives	**Endpoints**
and time away from work due to COVID-19	(WPAI) questionnaire • Change from baseline for health-related quality of life according to the SF-12 hybrid questionnaire
• Evaluate the impact of VIR-7831 versus placebo on the duration and the severity of COVID-19 clinical symptoms	• Mean change in FLU-PRO Plus Total score (AUC through Days 14, 21) and Domain scores (AUC through Days 7, 14, 21) comparing Vir-7831vs. Placebo • The proportion of participants with symptom alleviation using the FLU-PRO Plus at Days 7, 14, 21

1.2.2. Estimands

Each study objective is presented below with additional information, including prespecified estimands with related attributes.

Table 2 Estimands

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure
Primary Objective: Evaluate the efficacy of VIR-7831 versus placebo in preventing the progression of mild/moderate COVID-19 disease	Primary	Proportion of participants who have progression of COVID-19 through Day 29 as defined by: 1. Hospitalization > 24 hours for acute management of illness OR 2. Death	Entire trial population randomized (ITT)	Data analyzed as collected (treatment policy strategy)	Relative Risk
	Supplementary	Proportion of participants who	Entire trial population	Missing treated as	Relative Risk

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	**Analysis Set**	**Intercurrent Event Strategy**	**Population Level Summary Measure**
		have progression of COVID-19 through Day 29 as defined by: 1. Hospitalization > 24 hours for acute management of illness OR 2. Death	randomized (ITT)	progressions (composite strategy)	
	Supplementary	Proportion of participants who have progression of COVID-19 through Day 29 as defined by: 1. Hospitalization	Per-Protocol	Data analyzed as collected (treatment policy strategy)	Relative Risk

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	**Analysis Set**	**Intercurrent Event Strategy**	**Population Level Summary Measure**
		> 24 hours for acute management of illness OR 2. Death			
Secondary Objective: Evaluate the efficacy of VIR-7831 versus placebo in preventing the progression of mild/moderate COVID-19	Secondary	Proportion of subjects who have progression of COVID-19 as defined by visit to a hospital emergency room for management of illness, hospitalization >24 hrs or death at Day 8, Day 15, Day 22 or Day 29	ITT	Data analyzed as collected (treatment policy strategy)	Relative Risk

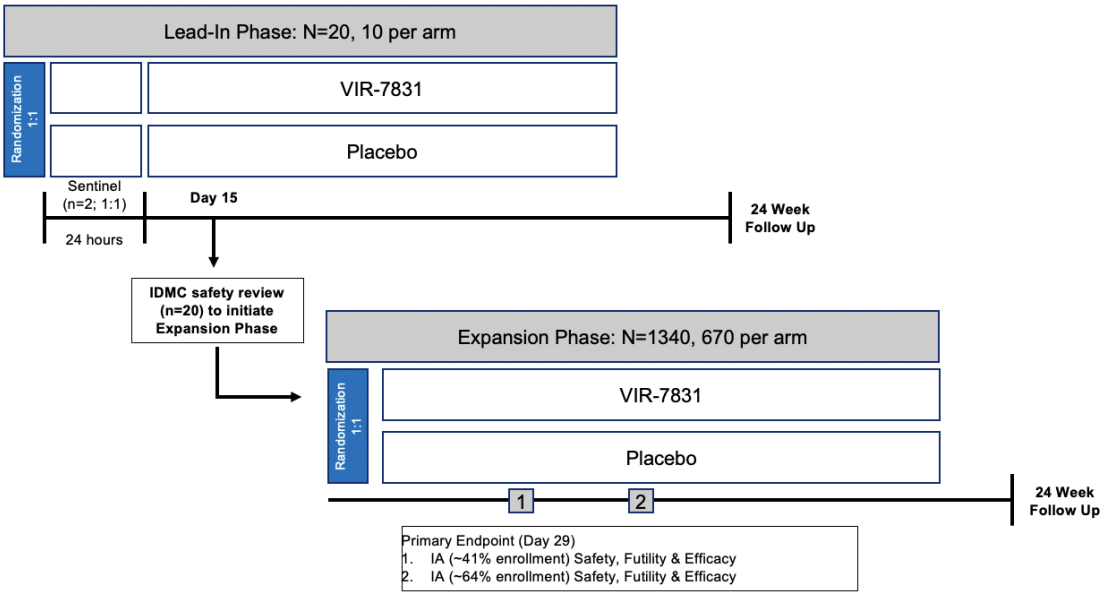
Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure
Secondary Objective: Evaluate the efficacy of VIR-7831 versus placebo in preventing COVID-19 respiratory disease progression	Secondary	Proportion of participants who progress to develop severe or critical respiratory COVID-19 as manifest by requirement for and method of supplemental oxygen at Day 8, Day 15, Day 22, or Day 29	ITT	Participants who die will be considered to have progressed to severe or critical respiratory COVID-19 by requirement for and method of supplemental oxygen For all other intercurrent events data will be analyzed as collected (treatment policy strategy)	Relative Risk
Secondary Objective: Evaluate the impact of VIR-7831 versus placebo in the duration and severity of COVID-19 clinical symptoms	Secondary	Mean change in FLU-PRO Plus total score comparing Vir-7831 vs. Placebo (AUC through Day 7)	ITT	Data analyzed as collected (treatment policy strategy)	Difference in Average
		Time to symptom alleviation using the FLU-PRO	ITT	Data analyzed as collected (treatment policy strategy)	Kaplan Meier Estimates

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure
		Plus			
Secondary Objective: Evaluate the efficacy of VIR-7831 versus placebo in reducing duration of SARS-CoV-2 viral shedding	Secondary	Change from baseline in viral load in nasal secretions by qRT-PCR at Day 8	ITT	Data analyzed as collected (treatment policy strategy)	Difference in Average
Secondary Objective: Evaluate the efficacy of VIR-7831 versus placebo in preventing mortality	Secondary	29-day, 60-day, and 90-day all-cause mortality	ITT	Data analyzed as collected (treatment policy strategy)	Kaplan Meier Estimates
Secondary Objective: Determine the safety and tolerability of VIR-7831 compared to placebo	Secondary	Occurrence of adverse events (AEs)	Safety	Data analyzed as collected (treatment policy strategy)	Frequency and Percentage of Participants
	Secondary	Occurrence of serious adverse events (SAEs)	Safety	Data analyzed as collected (treatment policy strategy)	Frequency and Percentage of Participants

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure
	Secondary	Occurrence of adverse events of special interest (AESI)	Safety	Data analyzed as collected (treatment policy strategy)	Frequency and Percentage of Participants

1.3. Study Design

Overview of Study Design and Key Features



Design Features	• This study is a randomized, double-blind, multi-center, placebo-controlled group-trial of VIR-7831, a monoclonal antibody (mAb) against SARS-CoV-2 for the prevention of progression of mild/moderate COVID-19 in high risk participants, with interim monitoring to allow early stopping for futility, efficacy, or safety. • The study will comprise two parts: lead-in phase and expansion phase. • The lead-in phase will enrol 20 non-hospitalized participants who have early, mild/moderate COVID-19 and are at high risk of disease progression. These participants will be given a single intravenous infusion of VIR-7831 and confined in the in-patient setting for 7 days for safety monitoring. • Following a safety assessment of unblinded data by an IDMC, the expansion phase will progress, where an additional 1340 participants with early, mild/moderate COVID-19 and who are at high risk of disease progression will be enrolled. • Unblinded safety and efficacy data will be reviewed regularly by an IDMC. • The first safety review during the expansion phase will include available safety and tolerability data though Day 15 from a total of 60 patients (30 per arm). If deemed necessary, an additional safety review will be conducted after approximately 100 additional participants are enrolled (50 per arm) to review safety and tolerability data through Day 15. IDMC members will also review all SAEs via draft CIOMS prior to each IDMC review

Overview of Study Design and Key Features	
	meeting. • Two interim analyses will be conducted at approximately 41% of participants enrolled (280 per arm) and approximately 64% of participants enrolled (435 per arm) to evaluate safety, futility, and efficacy based on data through Day 29 for these participants.
Study intervention	• VIR-7831 500mg or Placebo, single intravenous infusion • All participants will receive SoC as per institutional protocols, in addition to the study intervention.
Study intervention Assignment	• Participants with early, mild/moderate COVID-19 who are at high risk for progression of disease will be randomized 1:1 to receive a single, intravenous infusion of either VIR-7831 or equal volume saline placebo. • Participants will be stratified by age (≤70 vs. >70 years old)**,** symptom duration (≤3 days vs. 4-5 days) and region (North America, South America, Europe, Asia, Rest of World (RoW). • All participants will be centrally randomized using an Interactive Web Response System (IWRS). • Investigators will remain blinded to each participant's assigned study intervention throughout the course of the study. In order to maintain this blind, the study pharmacist will be responsible for the reconstitution and dispensation of all study intervention and will endeavor to ensure that there are no differences in time taken to dispense following randomization.
Interim Analyses	• An IDMC will actively monitor interim unblinded safety (Lead-In phase) and interim unblinded safety and efficacy data (Expansion phase) to make recommendations to Vir regarding ongoing study conduct as detailed in the IDMC Charter. • Interim analysis will be used to assess safety (due to harm), futility due to lack of efficacy and efficacy. Safety stopping rules will be put in place for the IDMC to stop the study if there is an excess risk of mortality or other events of interest. Full details of timing of analyses and all stopping criteria will be given in the IDMC charter. • Interim analyses to assess efficacy and futility are planned when approximately 41% and 64% of the required number of participants have reached Day 29 visit. • The analysis of the primary endpoint at the interim analyses will be the same as that described for the primary endpoint at the final Day 29 analysis. Group sequential techniques will be used to adjust stopping boundaries to reflect the actual number of participants with available data for primary endpoint at the time

Overview of Study Design and Key Features	
	of each interim analysis. • The full decision criteria for the futility due to lack of efficacy and efficacy rules along with additional information to be presented to the IDMC will be defined in the IDMC charter and associated documents, as well as the analysis plan. • The Joint Safety Review Team (JSRT) comprised of team members from clinical research, pharmacovigilance and statistics will review blinded safety data at regular intervals throughout the conduct of the study and determine if a safety concern based on instream blinded data review needs to be escalated to the IDMC. Details of the JSRT process is recorded in relevant SRT documents. • Following all participants completing the final follow-up visit at Week 24, a final safety analysis will be performed.

2. STATISTICAL HYPOTHESES / SUCCESS CRITERIA

The primary objective of this study is to evaluate the efficacy or VIR-7831 versus placebo in preventing the progression of COVID-19 disease.

The primary endpoint is proportion of participants who progression of COVID-19 as defined by hospitalization > 24 hours or death through Day 29.

If we denote the proportion of participants with progression of COVID-19 in the VIR-7831 arm as P_A, and the proportion of participants with progression of COVID-19 in the placebo arm as P_B, the following null and alternative hypothesis for superiority is as follows:

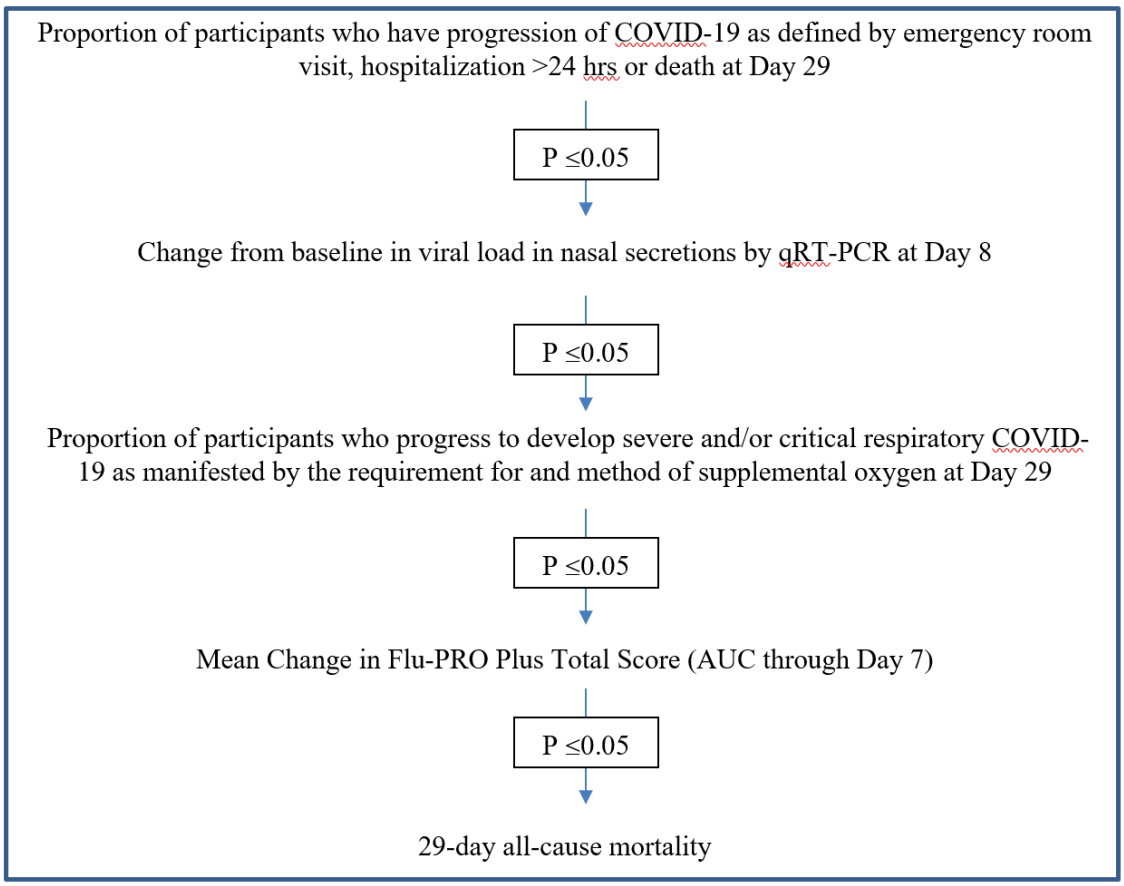
H0: $P_A = P_B$

H1: $P_A \neq P_B$

2.1. Multiple Comparisons and Multiplicity

The study will utilize a group sequential design with two interim analyses to asses both futility due to lack of efficacy and overwhelming efficacy. A Lan-DeMets [Lan, 1983] alpha-spending function to control the type I error for the primary endpoint will be used, using a Pocock analogue rule for futility and a Hwang-Shih-DeCani ($\gamma = 1$) analogue rule for efficacy [Hwang, 1990].

Secondary endpoints will only be formally analyzed at the final Day 29 analysis, secondary endpoints will be tested with alpha level of 5% (two-sided). The testing of secondary endpoints is adjusted for multiplicity by using the following hierarchy:



As such the full alpha is transferred down between each endpoint/hypothesis. All secondary endpoints in the testing hierarchy will present the nominal p-value of the analysis, and an overall summary of progress through each stage of the testing hierarchy will be provided.

3. SAMPLE SIZE DETERMINATION

The expected sample size is up to 680 participants per arm.

The study will utilize a group sequential design with two interim analyses to asses both futility due to lack of efficacy and overwhelming efficacy. A Lan-DeMets alpha-spending function to control the type I error will be used, using a Pocock analogue rule for futility and Hwang-Shih-DeCani ($\gamma = 1$) analogue rule for efficacy [Lan, 1983, Hwang, 1990].

Approximately 1360 (680 per arm) participants will be randomly assigned to study introversion. A total sample size of 1360 will provide approximately 90% power to detect a 37.5% relative efficacy in reducing progression of COVID-19 through Day 29 at the overall two-sided 5% significance level with assumed progression of COVID-19 rates of 16% in the placebo arm and 10% in the VIR-7831 arm, respectively. The minimal detectable efficacy for this design at the final efficacy analysis is approximately 25% if disease progression rates is 16% in the placebo arm.

4. ANALYSIS SETS

Analysis Set	Definition / Criteria	Analyses Evaluated
Screened	All participants who were screened for eligibility	Study Population
Enrolled	All participants who entered the study Note screening failures (who never passed screening even if rescreened) and participants screened are excluded from the Enrolled analysis set as they did not enter the study	Study Population
Intent-to-Treat	All participants who were randomly assigned to study intervention in the study. Data should be reported according to the randomized treatment	Study Population Efficacy
Safety	All participants who received at least one dose of study intervention. Data should be reported according to the actual treatment received	Study Population Safety
Per-Protocol (PP)	All participants in the ITT analysis set for whom there were no important protocol deviations that impact the primary analyses. Data should be reported according to the randomized treatment. Specific details of important protocol deviations that would exclude participants from the PP analysis set are provided in	Efficacy The PP set will generally not be used for analyses if this analysis set comprises more than 95% or less than 75% of the ITT analysis set.

Analysis Set	Definition / Criteria	Analyses Evaluated
	Section 4.1.1	
Pharmacokinetic (PK)	All participants in the Safety analysis set who had at least 1 non-missing PK assessment (Non-quantifiable [NQ] values will be considered as non-missing values). Data should be reported according to the actual treatment	PK
Virology	All participants in the ITT analysis set with a central lab confirmed quantifiable baseline nasopharyngeal swab. Data should be reported according to the randomized treatment.	Efficacy (Virology)

4.1. Protocol Deviations

Important protocol deviations are a subset of protocol deviations that may significantly impact the completeness, accuracy, and/or reliability of the study data or that may significantly affect a subject's rights, safety, or well-being. For example, important protocol deviations may include enrolling subjects in violation of key eligibility criteria designed to ensure a specific subject population or failing to collect data necessary to interpret primary endpoints, as this may compromise the scientific value of the trial.

Important protocol deviations (including deviations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarized.

Protocol deviations will be tracked by the study team throughout the conduct of the study. These protocol deviations will be reviewed to identify those considered as important as follows:

Data will be reviewed prior to unblinding and freezing the final Day 29 efficacy database to ensure all important deviations are captured and categorized in the protocol deviations SDTM dataset.

This dataset will be the basis for the summaries of important protocol deviations.

A separate summary of all inclusion/exclusion criteria deviations will also be provided. This summary will be based on data as recorded on the inclusion/exclusion page of the eCRF.

4.1.1. Definitions for Per Protocol Analysis Set

Important protocol deviations (including deviations related to study inclusion/exclusion criteria, conduct of the trial, patient management or patient assessment) will be summarized and listed.

All protocol deviations deemed important will result in exclusion from the Per Protocol analysis population. These will be reviewed in a blinded manner during the course of the study; any additional important deviations deemed to result in exclusion from the Per-Protocol analysis set will be documented prior to unblinding of the study (Refer to Protocol Deviation Management Plan).

5. STATISTICAL ANALYSES

5.1. General Considerations

5.1.1. General Methodology

The ITT Analysis Set will be used for all Study Population analyses and Efficacy analyses, unless otherwise specified. Safety Analysis Set will be used for selected Study Population summarises and all safety analyses. The PK Analysis Set will be used for all PK analyses.

In the case of a difference between the stratification assigned at the time of randomization and the data collected in the eCRF, the analyses will be performed based on the data collected in the eCRF, not the assigned stratum at randomization. All analyzes will be adjusted for duration of symptoms (≤3 days vs. 4-5 days), age (≤70 vs. >70 years old) and region (North America, South America, Europe, Asia, RoW) and also gender (male, female).

If there are convergence issues, duration of symptoms and age will be fitted as continuous rather than categorical covariates for primary and all secondary/exploratory analyses, as applicable. Should one randomisation region dominate the interim analysis then region may be considered for removal from the model to aid convergence issues.

Confidence intervals will use 95% confidence levels unless otherwise specified.

Unless otherwise specified, continuous data will be summarized using descriptive statistics: n, mean, standard deviation (std), median, minimum and maximum. Log-transformed data descriptive statistics will also present the geometric mean and coefficient of variation for data that is transformed back to the original scale. Categorical data will be summarized as the number and percentage of participants in each category.

Efficacy data will be reported according to nominal time of clinical visits. Hospitalization status, oxygen supplementation and healthcare resource utilization use will be slotted against the clinical visit date for summaries and analyses by visit.

All data will be reported according to the nominal time of clinic visits and assessments as specified in the protocol:

- Day 15, Day 18, Day 22, where data within ±1 days of the target day may be used if data is not recorded on the actual day
- Day 29, where data within ±2 days of the target day may be used if data is not recorded on the actual day
- Day 43, where data within ±3 days of the target day may be used if data is not recorded on the actual day
- Day 57, where data within ±4 days of the target day may be used if data is not recorded on the actual day
- Day 85, Day 113, Day 141, Day 169 where data within ±7 days of the target day may be used if data is not recorded on the actual day

Laboratory assessments will be analyzed based on the timepoint information recorded in the eCRF. Data for unscheduled visits, will be assigned to a scheduled visit if the visit date falls within the corresponding visit window as specified below.

5.1.2. Baseline Definition

For all endpoints the baseline value will be the latest pre-dose assessment with a non-missing value, including those from unscheduled visits and the screening visit. If time is not collected, Day 1 assessments are assumed to be taken prior to first dose and used as baseline.

Unless otherwise stated, if baseline data is missing no derivation will be performed and baseline will be set to missing.

5.1.3. Multicenter Studies

It is anticipated that patient accrual will be spread thinly across centers and summaries of data by center would unlikely be informative and will not, therefore, be provided.

Summaries by region (North America, South America, Europe, Asia, RoW as applicable).

5.1.4. Intercurrent Events

In general, the following may be considered intercurrent events/events leading to missing data if not part of the endpoint definition:

- Withdrawal of consent or lost-to-follow up
- Use of approved or authorized COVID-19 therapy
- Death
- Hospital Discharge
- Transfer to other facilities
- Admittance to ICU

The intercurrent event handling strategies are described in Table 2 Estimands and further expanded upon in Section 5.2 and Section 5.3.

In general, unless otherwise specified, the handling strategy for all these intercurrent events will be based on a treatment policy approach; specifically, the effects estimated will be based on initial randomized treatment arm regardless of whether the participant had experienced an intercurrent event. If possible, data will continue to be collected after the occurrence of the intercurrent event, until the participant either completes the study or withdraws from the study before completion.

5.1.5. Missing Data Handling Rules

Unless otherwise stated the following rules will be applied to all endpoints:

- If a participant experiences an intercurrent event but data continues to be collected, their data will be analyzed per their original randomized intervention.
- Missing data can occur due to study withdrawal or participants lost to follow-up before the completion of the study or due to intermittent missing values (i.e. data between two non-missing assessments).
- Missing data will be imputed under a missing at random (MAR) assumption using a multiple imputation (MI) model. The MI model will include covariates; treatment, duration of symptoms (≤3 days vs. 4-5 days), age group (≤70 vs. >70 years old), region (North America, South America, Europe, Asia and RoW), gender (male, female) and baseline of the variable of interest (if appropriate). More details will be provided in Appendix 3.

5.2. Primary Endpoint Analyses

5.2.1. Definition of Endpoint

The primary endpoint will be proportion of participants who have progression of COVID-19 as defined as hospitalization > 24 hours for acute management of illness (i.e. excludes hospitalization for observation only, hospitalization for elective surgeries and procedures, etc.) status or mortality status (death) through Day 29.

5.2.2. Main Analytical Approach

The primary endpoint will be summarized using counts and proportions of the number of participants who have progression of COVID-19 and will be analyzed using a log-binomial regression model adjusting for duration of symptoms (≤3 days vs. 4-5 days), age (≤70 vs. >70 years old), region (North America, South America, Europe, Asia and RoW) and gender (male, female).

The relative risk of progression will be calculated from the log-binomial generalised linear model. Should there be issues with model convergence then duration of symptoms and age (10 year increments) will be fitted as continuous covariates. Should one randomization region dominate the interim analysis then region may be considered for removal from the model to aid convergence issues. In the event the log-binomial model

still fails to converge, then an exact Poisson regression model with robust sandwich estimators will be utilized.

At each of the interim and final efficacy analysis, appropriate CI based on the adjusted significance level will also be provided. For example, if the remaining alpha for the final analysis is 0.024, 97.6% CI will be provided for the final efficacy analysis.

As patients may attend their Day 29 visit up to 2 days prior or following the planned date, the protocol defined visit windows will be used to define the endpoint. Further details on analysis windows will be provided in the Output and Programming Specification (OPS).

5.2.2.1. Decision Criteria

Guidelines for stopping the study for futility (*e.g*. lack of efficacy) or overall success (*e.g*. evidence of efficacy) based on the primary endpoint are detailed below.

Futility due to lack of efficacy and study success due to overwhelming benefit of treatment will formally be assessed 29 days after approximately 280/arm and 435/arm participants have been randomised and followed for 29 days.

Study stopping criteria have been defined using group sequential design methodology, using an alpha-spending function to control the type I error with two interim analyses for futility, using Pocock analogue rules; and two for efficacy, using the Hwang-Shih-DeCani ($\gamma = 1$) analogue rules [Hwang, 1990].

Table 3 gives the stopping boundaries based on p-values (one-sided) from the planned interim analysis and on the Z score scale. The p-values or Z scores at each interim analysis will be plotted against the boundaries and if either the futility or efficacy success boundary is crossed the study should be recommended to stop.

The boundaries are based on the estimated amount of information at each interim. If the amount included in the interim analysis differs from these values, the actual boundaries will be re-determined based on the exact amount of information at the time of the interim analyses. Exact amount of information will be determined from the number of participants randomised at 29 days prior to data cut.

Boundaries will be determined using PASS 2019, using the group sequential tests for two proportions (simulation procedure) with the following criteria: 2-sided overall 5% significance level (symmetric), equal allocation, 3 planned stages at actual and planned (future) information proportions (minimum 5 decimal places), no continuity correction, 0.1 zero count adjustment added to zero cells only and non-binding futility boundaries and hold-out efficacy boundaries.

Table 3 Planned Formal Stopping Rules for Efficacy Success and Efficacy Futility

Stage/Formal Analysis	Target Sample Size to D29	Efficacy Success Boundaries				Efficacy Futility Boundaries			
		p-value		Z-Score		p-value		Z-score	
		Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
1	560	<0.01335	<0.01335	-2.2159	2.2159	>0.21147	>0.21147	-0.8013	0.8013
2	870	<0.00973	<0.00973	-2.3367	2.3367	>0.09363	>0.09363	-1.3187	1.3187
Final (D29)	1360	<0.01103	<0.01103	-2.2894	2.2894	>0.01103	>0.01103	-2.2894	2.2894

5.2.3. Sensitivity Analyses

As a sensitivity to the missing data, a tipping point analyses will be performed.

The underlying response rate among those subjects with missing response status in each arm will be tested ranging between 0 and 1. This analysis will be two-dimensional, i.e. will allow for assumptions about the assumed response rate (and thus missing outcomes) in the two arms to vary independently, furthermore they will include scenarios where dropouts on VIR-7831 have worse outcomes than dropouts on placebo. For combinations of the assumed response rates in the two arms, the number of additional responders among subjects with missing response will be imputed multiple times by drawing from a binomial distribution. The risk ratio and associated standard error for each imputed dataset will be calculated and results combined using Rubin's rules to calculate the test statistic and the corresponding p-value. Results will be presented via a heatmap.

5.2.4. Supplementary Analyses

A supplementary analysis for the primary endpoint will also be conducted on the Per-Protocol analysis set, if this comprises more than 95% or less than 75% of the ITT analysis set.

As a supplementary analysis, a composite estimand, where withdrawals will be treated as having progressed, will be presented as a supplementary estimand.

Further estimand strategies may be investigated and will be described in a separate technical document.

5.2.5. Subgroup Analyses

Subgroup analyses of the duration of symptoms (≤3 days vs. 4-5 days), age (≤70 vs. >70 years old), gender and region will be performed using the same methodology as the

primary endpoint. Results will be presented as per primary analyses and, in addition, a forest plot of the relative risk ratios and the corresponding 95% CIs will be generated.

If there are insufficient numbers of participants in each category to be able to perform the analyses, then categories may be combined, or the analyses may not be performed if there is not an appropriate grouping.

Additional subgroup analyses may be performed and will be described in a separate technical document.

5.3. Secondary Endpoints Analyses

Secondary endpoints will be summarized for the IDMC reviews, and only analyzed at the final Day 29 analysis. The estimands are described in Table 2.

Confirmatory secondary endpoints are those secondary endpoints for which a label claim is pursued as part of the confirmatory hypotheses for which the type 1 error is controlled (multiplicity adjustment).

5.3.1. Definition of Endpoints

5.3.1.1. Safety and tolerability

Adverse events analyses including the analysis of adverse events (AEs), Serious AEs (SAEs) and other significant AEs will be based on GSK Core Data Standards.

An overview summary of AEs, including counts and percentages of participants with

- any AE
- any SAE

will be provided as outlined in Section 5.5.2.

Since it will not be possible to delineate in a single participant whether the hospitalization is directly related to COVID-19 complications or could be related to VIR-7831 causing more severe disease due to ADE, all hospitalizations regardless of cause will be included in the primary endpoint and will also be counted as serious adverse events. To inform on the number and nature of non-COVID-19 adverse events and serious adverse events, additional safety summaries will be performed in which select, pre-specified terms consistent with known progression of COVID-19 disease will be excluded. The MedDRA terms to be excluded will be pre-defined and documented in the study OPS.

5.3.1.2. Immunogenicity of VIR-7831

The incidence and titers (if applicable) of serum ADA to VIR-7831 will be listed and summarized for the Week 24 safety summary.

Additional immunologic analyses may be conducted and will be documented in a separate technical document.

5.3.1.3. Pharmacokinetics of VIR-7831

Serum pharmacokinetic concentrations and parameters will be listed and summarized by visit.

Pharmacokinetic parameters will be calculated by standard non-compartmental analysis according to current working practices and using the currently supported version of WinNonlin. All calculations of non-compartmental parameters will be based on actual sampling times, and will be listed and summarized using descriptive statistics. PK parameters may include, but not be limited to C_{max}, C_{last}, C_{D29}, t_{max}, t_{last}, AUC_{inf}, AUC_{last}, $AUC_{D1\text{-}D29,}$ $\%AUC_{exp}$, $t_{1/2}$, Vz, Vss, CL. Inclusion of PK parameters in summaries will be at the discretion of the PK scientist. Any parameters not included in summaries should be flagged in the individual listings with an explanation for the exclusion.

5.3.1.4. Proportion of Participants Who Have Progression of COVID-19

Defined as per primary endpoint, see Section 5.2.1 plus additional emergency room visits (inpatient or out-patient visits of any duration).

5.3.1.5. Proportion of Participants Who Progress to Develop Severe or Critical Respiratory COVID-19

Participants are defined as progressing to develop severe respiratory COVID-19 if they require supplemental oxygen either by nasal cannula, face mask, high-flow oxygen devices, or non-invasive ventilation.

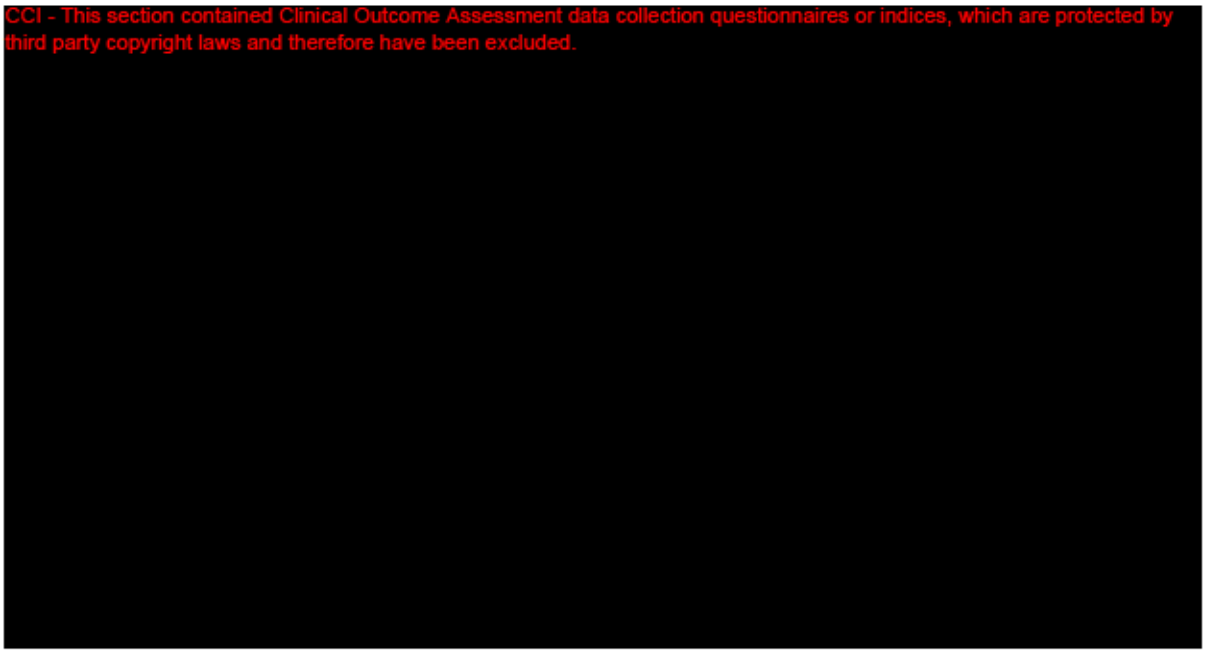
Participants are defined as progression to critical respiratory COVID-19 if they require invasive mechanical ventilation or ECMO.

Participants who die prior to Day 29 without first having received supplemental oxygen will be considered to have met the endpoint (composite estimands strategy).

Further detail on the severity of respiratory COVID-19 based on the proportion of participants meeting each tier of the Vir modified version of the NIAID Ordinal Scale for Clinical Improvement (below) will be summarized.

Table 4 Ordinal Scale for Clinical Improvement

CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.



5.3.1.6. Severity and Duration of COVID-19 Clinical Symptoms

Severity and duration of COVID-19 clinical symptoms of COVID-19-related illness will be determined from averaged change (AUC as determined by the trapezoidal rule) though 7 days in the total and domain scores (nose, throat, eyes, chest/respiratory, gastrointestinal and body/systemic) from the FLU-PRO Plus patient-reported outcome instrument and will also combine additional scores for smell/taste.

In addition, time-to sustained (>=48 hours) symptom alleviation as measured by the FLU-PRO Plus over first 21 days will be compared. The definition of sustained symptom alleviation is included in Appendix 4.

COVID-19 Clinical Symptoms analyses will be provided for the population, and in the subgroup of participants who reported at least two moderate symptoms at baseline.

5.3.1.7. Reduction in SARS-CoV-2 Viral Load.

Reduction of SARS-CoV2 viral shedding will be determined by the change from baseline of SARS-CoV2 nasal viral load at Day 8. Viral load will be log10-transformed prior to summary and analysis.

5.3.1.8. All-cause Mortality

Participants will meet the endpoint if they have died due to any cause prior to Day 29, Day 60 and Day 90. Day 60 and Day 90 all-cause mortality will be presented in the Week 24 Safety Analysis.

5.3.2. Main Analytical Approach

5.3.2.1. Binary Endpoints

The binary endpoints:

- Participants who have progression of COVID-19
- Participants who progress to develop severe or critical respiratory COVID-19

Binary secondary endpoints will be analysed as per Section 5.2.2

5.3.2.2. Continuous Endpoints

The following secondary endpoints will be analysed as continuous endpoints:

- Severity of participant-reported symptoms of COVID-19-related illness using Flu-PRO Plus (Average Change from Baseline)
- Reduction of SARS-CoV-2 viral shedding using qRT-PCR

Flu-PRO Plus (AUC) will be analysed using analysis of covariance adjusted for baseline value, age group, time to onset of symptoms, gender and region at each visit.

Possible intercurrent events are withdrawal of consent and lost-to-follow-up on or before Day 29. The strategy for analysis for intercurrent events, will be to use the treatment policy strategy, with data analyzed as collected.

It is expected that missing data questions within the Flu-PRO Plus will be limited to paper PROs due to the set-up of ePRO devices. The scoring of the Flu-PRO Plus domain scores limits the extent of missed questions permitted when determining a domain score; the total score requires all domain scores to be calculated prior to determining the total score.

Thus, it is expected missing data will be due to incomplete paper PROs or lost to follow up data after discharge, and this will be imputed using a modified last-observation carried forward (mLOCF) approach for the final assessment only, due to the assumption incomplete paper PROs are not missing at random. That is if the Day 7 Flu-PRO Plus score is missing or incomplete, the last non-missing value will be carried forward for the calculation of average change from baseline through to Day 7; no other missing scores would be imputed.

SARS-CoV-2 viral load will be analysed using a mixed model repeated measures (MMRM) adjusted for baseline value, baseline value by visit, age group, time to onset of symptoms, gender and region.

It is expected missing data for SARS-CoV-2 qRT-PCR results will be due to invalid qRT-PCT test, or lost to follow up data after discharge, and thus are assumed missing at random within the analysis.

5.3.2.3. Time to Event Endpoints

The following secondary endpoints will be analyzed as time to event endpoints:

- All-cause mortality
- Sustained (>=48h) symptom alleviation as measured by adapted Flu-PRO Plus

It is expected that missing data questions within the Flu-PRO Plus will be limited to paper PROs due to the set-up of ePRO devices. Incomplete paper PROs are assumed to be missing not at random, thus symptom alleviation will use a LOCF approach for missing data. That is, questions will first be checked for imputation of missing questions using LOCF when determining whether symptom alleviation has been achieved; if a symptom alleviation cannot be determined missing symptom alleviation (yes/no) will be imputed using LOCF.

Time to event endpoints will be analyzed using nonparametric survival analysis methods, using Kaplan-Meier methods for summarizing data and analyzed using the log-rank test.

5.3.3. Sensitivity Analyses

Sensitivity analyses as per Section 5.2.3 may be performed on the binary endpoints.

5.3.4. Supplementary Analyses

Other estimands strategies may be investigated and will be described in a separate technical document.

5.3.5. Subgroup Analyses

Subgroup analyses will be performed on the secondary endpoints detailed in Section 2.1 per Section 5.2.5 and will use the methods described above.

5.4. Exploratory Endpoints Analyses

5.4.1. Definition of Endpoints

5.4.1.1. Total Number of Ventilator Days

A participant is defined as being on a ventilator if have received invasive mechanical ventilation or ECMO as a form of oxygen therapy. A participant may have more than one period of ventilation during the analysis time-point follow-up included in the total number of ventilator days.

Participants who die prior to Day 29/Day 90 without being on a ventilator will be considered to have met the endpoint (composite estimands strategy).

5.4.1.2. Total Intensive Care Length of Stay

A participant is defined as requiring intensive care if they received an intensive care unit in-patient healthcare encounter of any duration. A participant may have more than one period of intensive care stays included in the total intensive care length of stay.

Participants who die prior to Day 29/Day 90 without being admitted to intensive care will be considered to have met the endpoint (composite estimands strategy).

5.4.1.3. Total Hospital Length of Stay

A participant is defined as requiring hospital stay if they received any in-patient healthcare (A&E, general ward or ICU) encounter of any duration. A participant may have more than one period of hospital stay included in the total hospital length of stay.

Participants who die prior to Day 29/Day 90 without being admitted to hospital will be considered to have met the endpoint (composite estimands strategy).

5.4.1.4. Reduction in the Duration of SARS-CoV-2 Viral Shedding.

Duration of SARS-CoV2 viral shedding will be determined by the time to obtain a confirmed negative PCR test for SARS-CoV-2 by Day 29. A confirmed negative PCR is defined as first of two or more consecutive negative (no SARS-CoV2 detected) PCR tests. If the first negative PCR occurs at Day 29, this will be regarded as confirmed for analysis purposes.

Reduction of SARS-CoV2 viral shedding (nasal secretions and blood) will be determined by the change from baseline of SARS-CoV2 obtained in a PCR test at each visit. Time weighted average change from baseline though the first 7 days will also be presented. Viral load will be log10-transformed for summary and analysis purposes.

5.4.1.5. Proportion of Participants Hospitalized Due to Non-Respiratory Complications of COVID-19

A participant is defined as hospitalized due to non-respiratory complications of COVID-19 for if the primary reason was for hospitalization was due to cardiac, renal, neurologic, hematologic events.

Participants who die prior to Day 29 without being admitted to hospital will be considered to have met the endpoint (composite estimand strategy).

5.4.1.6. On-treatment Emergence of SARS-CoV-2 Resistant Mutants Against VIR-7831

Summaries of viral resistance mutations will be provided based on frequency counts.

5.4.1.7. Host Immune Responses and Exploratory Biomarkers

Further details on genetic, cellular, transcriptomic, and proteomic parameters will be provided in a supplemental analysis plan.

5.4.1.8. Severity and Duration of COVID-19 Clinical Symptoms

Severity and duration of COVID-19 clinical symptoms of COVID-19-related illness will be determined from averaged change (AUC as determined by the trapezoidal rule) though 14 and 21 days in the total and domain scores (nose, throat, eyes, chest/respiratory, gastrointestinal and body/systemic) from the Flu-PRO Plus patient-reported outcome instrument and will also combine additional scores for smell/taste.

In addition, the proportion of participants achieving sustained (>=48 hours) symptom alleviation as measured by the FLU-PRO Plus will be presented. The definition of sustained symptom alleviation is included in Appendix 4.

COVID-19 Clinical Symptoms will be provided for the population, and in the subgroup of participants who reported at least two moderate symptoms as baseline.

5.4.1.9. Work Productivity and Activity Impairment (WPAI)

At Baseline, participants were queried as to whether they were able to socially distance and either work on-site or remotely.

Mean change in the WPAI scores for absenteeism (work time missed), presenteeism (impairment at work / reduced on-the-job effectiveness), work productivity loss (overall work impairment / absenteeism plus presenteeism) and activity impairment will be assessed at each visit.

5.4.1.10. Health-Related Quality of Life According to the SF-12 Hybrid

Mean change in the Health Status SF-12 plus the full SF-36 domains of vitality and physical role will be completed. The total score, the physical component summary (PCS), mental component summary (MCS) and the SF-36 domain scores for vitality and physical role will be assessed at each visit.

5.4.2. Main Analytical Approach

5.4.2.1. Binary Endpoints

The following exploratory endpoints will be analysed as continuous endpoints:

- Proportion of participants hospitalized due to non-respiratory complications of COVID-19
- Proportion of participants with sustained (>= 48 hours) symptom alleviation

Binary secondary endpoints will be analysed as per Section 5.2.2.

The strategy for missing data in derivation of symptom alleviation based on Flu-PRO Plus is as described in Section 5.3.2.3.

5.4.2.2. Continuous Endpoints

The following exploratory endpoints will be analysed as continuous endpoints:

- Total number of ventilator days from randomization through 90 days
- Total intensive care length of stay
- Total hospital length of stay
- Reduction in SARS-CoV2 viral load (nasal secretions and blood)
- Severity and Duration of COVID-19 Clinical Symptoms as measured by Flu-PRO Plus.
- Work Productivity and Activity Impairment (WPAI)
- Health-related quality of life according to the SF-12 hybrid

SARS-CoV-2 viral load will be analysed using a mixed model repeated measures (MMRM) adjusted for baseline value, baseline value by visit, age group, time to onset of symptoms, gender and region as described in Section 5.3.2.2.

All other continuous endpoints will be analyzed using analysis of covariance adjusted for age group, time to onset of symptoms, region and gender at each visit. Additionally, for Flu-PRO Plus and SF-12 the baseline measurement will be included.

The strategy for handling intercurrent events and missing data for SARS-CoV-2 viral load and Flu-PRO Plus is as described in Section 5.3.2.2.

Continuous endpoints will be summarised using the trimmed means approach (Permutt, 2017) where the proportion of data to be trimmed will be determined by the amount of missing data due to death or lost-to-follow up/consent withdrawn. For hospital and ICU/Ventilation use, the trimmed means will only be utilized if at least 10% of subjects overall have the event of interest.

Descriptive statistics (n, arithmetic mean, standard deviation [SD] standard error [SE], 95% CI, minimum, median and maximum) will be calculated by treatment. Total number of ventilator days, total intensive care length of stay and total hospital length of stay will only be analyzed if at least 10% of participants have one or more recorded event of interest.

5.4.2.3. Time-to-event endpoints

The following exploratory endpoints will be analyzed as time to event endpoints:

- Time to undetectable nasal SARS-CoV2 viral load

Time to event endpoints will be analyzed using nonparametric survival analysis methods, using Kaplan-Meier methods for summarizing data and analyzed using the log-rank test.

5.5. Safety Analyses

The safety analyses will be based on the “Safety” Analysis Set, unless otherwise specified.

5.5.1. Extent of Exposure

As this is a single dose study, summaries of exposure will be limited to the number of participants exposed and the number of participants with Interruptions or Infusion stopped early and not completed.

This will be presented with Subject Disposition.

5.5.2. Adverse Events

Adverse events analyses including the analysis of adverse events (AEs), Serious AEs (SAEs) and other significant AEs will be based on GSK Core Data Standards.

An overview summary of AEs will be produced, including counts and percentages of participants with:

- AEs
- AEs related to study intervention
- AEs leading to permanent discontinuation of study intervention,
- AEs leading to temporary interruption of study intervention
- Grade 3 and 4 AEs
- Grade 3 and 4 AEs related to study intervention,
- Grade 3 and 4 AEs leading to permanent discontinuation of study intervention,
- Grade 3 and 4 AEs leading to temporary interruption of study intervention
- SAEs,
- SAEs related to study intervention,
- Fatal SAEs
- Fatal SAEs related to study intervention will be produced.

In addition, an overall summary of safety criteria will be produced, including counts and percentages of participants with:

- All-cause mortality
- SAEs
- Renal events
- Cardiac events
- Pulmonary events

Adverse events will be coded using the standard Medical Dictionary for Regulatory Affairs (MedDRA dictionary) and graded by the investigator according to the DAIDS 2017 v2.1.

The frequency and percentage of AEs (all grades) will be summarized and displayed in in descending order by System Organ Class (SOC) and Preferred Term (PT). In the SOC

row, the number of participants with multiple events under the same SOC will be counted once.

A summary of number and percentage of participants with any adverse events by maximum grade will also be produced. AEs will be sorted by PT in descending order. The summary will use the following algorithms for counting the participant:

- **Preferred term row**: Participants experiencing the same AE preferred term several times with different grades will only be counted once with the maximum grade.
- **Any event row**: Each participant with at least one adverse event will be counted only once at the maximum grade no matter how many events they have.

A separate summary will be provided for study intervention-related AEs (all grades). A study intervention-related AE is defined as an AE for which the investigator classifies the possible relationship to study intervention as "Yes". A worst-case scenario approach will be taken to handle missing relatedness data, i.e. the summary table will include events with the relationship to study intervention as 'Yes' or missing. The summary table will be displayed by SOC and PT. Similarly, a summary of study intervention-related AEs by maximum grade will also be produced. The summary table will be displayed by PT only.

Other summaries of AEs to be produced will include:

- AEs/study intervention related AEs leading to permanent discontinuation from study,
- Most common AEs/ study intervention related AEs/ Grade 3 to 4 AEs/ study intervention related Grade 3 to 4 AEs, with common AE defined as any AE (PT) with an incidence of at least 5% in any of the treatment group
- All AEs excluding events related to COVID-19
- All AEs for subjects who received approved COVID-19 vaccination

All SAEs will be tabulated based on the number and percentage of participants who experienced the event. The summary tables will be displayed by SOC and PT. Separate summaries will also be provided for study intervention related SAEs. The summary tables will be displayed by PT.

A study intervention-related SAE is defined as an SAE for which the investigator classifies the relationship to study intervention as "Yes". A worst-case scenario approach will be taken to handle missing data, i.e. the summary table will include events with the relationship to study intervention as 'Yes' or missing.

Relative risks based on observed frequencies, for the proportions of participants with common (>=5% in either treatment group) AEs will be calculated for VIR-7831 versus placebo.

Relative risks with Exact unconditional confidence intervals will be presented graphically. Exact confidence limits will be computed by inverting two separate one

sided exact tests that are based on score statistic. Relative risks will not be calculated if there are no events in either of the two treatment arms being compared. The relative risks and exact confidence interval will be plotted on log base 10 scale.

The adverse events with calculated relative risks will be presented in decreasing order.

The full list of planned AE/SAE summaries will be as described in Section 4 of the study OPS.

5.5.2.1. Adverse Events of Special Interest

The following will be considered adverse events of special interest (AESI) for the purpose of analyses:

- Infusion related reactions (IRR) including hypersensitivity reactions; reactions on same day as infusion
- Immunogenicity related adverse drug reactions
- Adverse events potentially related to antibody dependent enhancement of disease

The summary of event characteristics will be provided for each AESI respectively, including:

- number of participants with any event,
- number of events,
- number of Grade 3/4 events,
- number of participants with any event that is serious,
- number of participants with any event that is related to study intervention,
- number of occurrences (One, Two, Three or more),
- maximum grade,
- maximum grade for events related to study intervention,
- outcomes and the action taken for the event.

The percentage will be calculated in two ways, one with number of participants with event as the denominator and the other with total number of participants as the denominator. The worst-case approach will be applied at participant level for the maximum grade, i.e. a participant will only be counted once as the worst case from all the events experienced by the participant. For action taken to an event, a participant will be counted once under each action, e.g. if a participant has an event leading to both study intervention discontinuation and dose reduction, the participants will be counted once under both actions.

The time to onset and duration of first event will also be summarized descriptively.

5.5.3. Additional Safety Assessments

5.5.3.1. Laboratory Data

Summaries of worst-case grade increase from baseline grade will be provided for all the lab tests that are gradable by DIAIDs 2017 v2.1. These summaries will display the number and percentage of subjects with a maximum post-baseline grade increasing from their baseline grade. Any increase in grade from baseline will be summarized along with any increase to a maximum grade of 2, any increase to a maximum grade of 3 and any increase to a maximum grade of 4. Missing baseline grade will be assumed as grade 0. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia separately.

For lab tests that are not gradable by DAIDS 2017 v2.1, summaries of worst-case changes from baseline with respect to normal range will be generated. Decreases to low, changes to normal or no changes from baseline, and increases to high will be summarized for the worst-case post-baseline. If a subject has a decrease to low and an increase to high during the same time interval, then the subject is counted in both the “Decrease to Low” categories and the “Increase to High” categories.

Separate summary tables for haematology, and chemistry laboratory tests will be produced. Liver function laboratory tests will be included with chemistry lab tests.

Summaries of hepatobiliary laboratory events including possible Hy’s law cases will be provided in addition to what has been described above. Possible Hy’s law cases are defined as any elevated alanine aminotransferase (ALT)>3×upper limit of normal (ULN), total bilirubin≥2×ULN and alkaline phosphatase (ALP)<3×ULN/missing. Total bilirubin≥2×ULN can be within 28 days following the ALT elevation and if direct bilirubin is available on the same day, it must be ≥ 35% of total bilirubin. ALP<3×ULN/missing means it is satisfied unless the ALP is ≥3xULN at the time of bilirubin elevation. The summary will be produced for worst case post baseline only. A plot for maximum post-baseline Total Bilirubin against ALT will also be produced, and will be repeated with the values standardized with respect to baseline, instead of ULN. This is to account for the individual variation at baseline

5.5.3.2. Vital Signs

Summaries of grade increase in systolic blood pressure (SBP) and diastolic blood pressure (DBP) will be provided separately. These summaries will display the number and percentage of participants with any grade increase, increase to Grade 2 and increase to Grade 3 for worst case post-baseline only. The grade definition for SBP is: Grade 0 (<120), Grade 1 (120-139), Grade 2 (140-159), Grade 3 (>=160). The grade definition for DBP is: Grade 0 (<80), Grade 1 (80-89), Grade 2 (90-99), Grade 3 (>=100). The summaries will be produced for worst case post baseline only.

In addition, summaries of respiratory rate, heart rate and temperature will be provided.

5.5.3.3. Oxygen Saturation

Summaries of actual and changes in blood oxygen saturation ($Sp0_2$) and proportion of subjects administered oxygen at each visit will be provided separately. Summaries of worst-case actual changes in FiO_2 and $Sp0_2/Fi0_2$ ratios will be summarised for those administered supplemental oxygen. Where FiO_2 is not collected, this will be estimated from the oxygen flow rate and device providing supplemental oxygen.

For participants administered supplemental oxygen, additional summary tables will display the number and percentage of participants with $\%SpO_2 < 92$ and/or $SpO_2/FiO_2 < 315$ at each visit and worst-case post-baseline.

Changes in requirement for respiratory support according to the following categories:

1. Room air
2. Low flow nasal canulae/face mask
3. Non-re-breather mask or high flow nasal cannulae/ non-invasive ventilation (including continuous positive airway pressure support)
4. Mechanical ventilation / extra-corporeal membrane oxygenation

Changes in requirement for respiratory support will be summarized using counts and proportions of the number of participants in each category.

In addition, a shift table from baseline to the level of respiratory support will be produced and the proportion of participants in each category will be summarised at each visit using a stacked bar chart. Individual subject profiles showing the time course of respiratory support will also be produced.

5.5.3.4. ECG

A summary of the number and percentage of participants with ECG findings will be summarized by treatment. The ECG findings to be summarized are the ECG interpretation, clinical significance of abnormal ECGs, and whether there was a clinically significant change from baseline. Participants with missing baseline values will be excluded from this summary.

5.6. Other Analyses

5.6.1. Population Pharmacokinetics

Sparse PK samples will be analysed using population PK approach. Parameters may include, but not be limited to C_{max}, C_{last}, t_{max}, t_{last}, AUC_{inf}, AUC_{last}, $AUC_{0\text{-}D28,}$ $\%AUC_{exp}$, $t_{1/2}$, Vz/F, CL/F, and will be listed and summarized using descriptive statistics.

A two-compartment IV bolus population pharmacokinetic model will be used to describe VIR-7831 pharmacokinetics. The model will fit concentration data in the log-domain with additive residual and takes the general form:

Log(CONC) = LNDV = log(Dose) + log(A exp(-ALPHA*TIME) + B*exp(-BETA*TIME)) + eps

Since infusion duration is very short compared to both distribution, ALPHA- and elimination BETA-phases, the bolus assumption is valid for monoclonal antibody administration. A physiological parameterisation of A, B, ALPHA and BETA will be used in terms of clearance (CL), inter-tissue peripheral clearance (Q), central volume (V1) and peripheral volume (V1). These physiological parameters will be related to bodyweight using conventional allometry (standardised to 70 kg), with fixed bodyweight powers of 0.75 for clearance and 1.00 for volume and subject to additional random effects (b1, b2). Observed sample times will be used for analysis and units will be days. Should convergence not be achieved by default Gaussian quadrature, the firo first-order method will be implemented instead.

5.7. Interim Analyses

Interim analysis will be used to assess safety (due to harm), futility due to lack of efficacy and efficacy. Safety stopping rules will be put in place for the IDMC to stop the study if there is an excess risk of harm (see IDMC Charter for further details.)

Interim analyses to assess efficacy and futility are planned when approximately 41% and 64% of the required number of participants have reached Day 29 visit. The analysis of the primary endpoint at the interim analyses will be the same as that described for the primary endpoint. Group sequential techniques will be used to adjust stopping boundaries to reflect the actual number of participants with available data for primary endpoint at the time of each interim analysis. The timing of the efficacy interim analyses and stopping rules are provided in Section 5.2.2.1.

Details of the interim analyses and required outputs are provided in the IDMC Charter and IDMC Analyses Plan.

The Joint Safety Review Team (JSRT) comprising physicians and scientists from Vir and GSK will review blinded safety data at regular intervals throughout the conduct of the study and determine if a safety concern based on instream blinded data review needs to be escalated to the IDMC. Details of the JSRT process is recorded in relevant SRT documents.

Following all participants completing the final follow-up visit at Week 24 a final safety analysis will be performed.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 Abbreviations and Trademarks

6.1.1. List of Abbreviations

Abbreviation	Description
AE	Adverse Event
A&E	Accident and Emergency
ADA	Antidrug Antibosy
AESI	Adverse Event of Special Interest
ALT	Alanine Aminotransferase
APX	Analysis Population Exclusion
AUC	Area Under the Curve
AST	Aspartate Aminotransferase
BiPAP	Bilevel Positive Airway Pressure
CI	Confidence Interval
CMH	Cochran–Mantel–Haenszel
COVID-19	Corona Virus Disease 2019
CPAP	Continuous Positive Airway Pressure
CV_b / CV_w	Coefficient of Variation (Between) / Coefficient of Variation (Within)
DBF	Database Freeze
DBP	Diastolic Blood Pressure
DBR	Database Release
ECG	Electrocardiogram
ECMO	Extracorporeal membrane oxygenation
eCRF	Electronic Case Record Form
FiO2	Concentration of Inspired Oxygen
Flu-PRO	Influenza-Patient Reported Outcome
GSK	GlaxoSmithKline
ICH	International Conference on Harmonization
ICU	Intensive Care Unit
IDMC	Independent Data Monitoring Committee
IND	Investigational New Drug
ITT	Intent-To-Treat
IWRS	Interactive Web-based Randomisation System
JSRT	Joint Safety Review Team
LOS	Length of Stay
mAb	monoclonal antibody
MAR	Missing at Random
MCMC	Markov Chain Monte Carlo
MCS	Mental Component Score
MI	Mutliple Imputation

Abbreviation	Description
NIAID	National Institute of Allergy and Infectious Diseases
OPS	Output and Programming Specification
PCI	Potential Clinical Importance
PCR	Polymerase chain reaction
PCS	Physical Component Score
PD	Protocol Deviation
PK	Pharmacokinetic
PP	Per Protocol
qRT	Quantitative Reverse Transcription
RoW	Rest of World
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SARS-CoV2	Severe Acute Respiratory Syndrome Coronavirus 2
SBP	Systolic Blood Pressure
SD/SE	Standard Deviation / Standard Error
SF-12	12-item short form survey
Sp02	Blood Oxygen Saturation
ULN	Upper Limit Normal
WPAI	Work Productivity and Activity Impairment

6.1.2. Trademarks

Trademarks of the GlaxoSmithKline Group of Companies
None

Trademarks not owned by the GlaxoSmithKline Group of Companies
None

6.2. Appendix 2: Changes to Protocol-Planned Analyses

There are no changes to the protocol-planned analyses.

6.3. Appendix 3: Statistical Modelling

6.3.1. Multiple Imputation

Multiple imputation will be utilized to impute data that is missing following withdrawal of the subject from the study. Each data type will use one set of imputations with all endpoints that depend on that data type being derived from the intermediate dataset.

The statistical model used for the multiple imputation data generation will use one of two approaches depending on data type:

- Continuous endpoints will use the Markov Chain Monte Carlo (MCMC) method with adjustment for covariates.
- Binary and ordinal endpoints will use a two-stage approach:
    - 1. First intermittent missing data will be imputed using MCMC followed by adaptive rounding [Carpenter, 2012] step to create a monotone structure.
    - 2. Second a monotone logistic regression imputation will be performed to impute data post-study withdrawal, in addition to the analyses covariates this model will include a parameter for COVID-19 progression (Progressed/Not Progressed). If logistic regression step fails to converge the following back-up options including:
    - i. A discriminant function [Brand, 1999] may be used in place of logistic.
    - ii. MCMC with adaptive rounding.

A sufficient number of imputations will be performed to ensure the stability of the estimates. The results of analysis from each complete imputed dataset will be combined using Rubin's rule.

Table 5 details the models, covariates and the seed to be used for the analyses.

Table 5 Multiple Imputation Specifications

Endpoint	Model	Covariates	Post-Baseline Timepoints Included in Prediction Model	Initial Seed
Proportion of participants who have progression of COVID-19 at Day 8, Day 15, Day 22, or Day 29	Monotone Binary Logistic Regression	Treatment, Age Group at Baseline, Duration of Symptoms	Day 8, Day 15, Day 22, or Day 29	6862
Proportion of participants who progress to	Monotone Binary Logistic Regression	Treatment, Age Group at Baseline,	Day 8, Day 15, Day 22, or Day 29	376

Endpoint	Model	Covariates	Post-Baseline Timepoints Included in Prediction Model	Initial Seed
develop severe or critical respiratory COVID-19		Duration of Symptoms		
29-day, 60-day, and 90-day all-cause mortality	Monotone Binary Logistic Regression	Treatment, Age Group at Baseline, Duration of Symptoms	Day 29, 60 & 90	7371

The seeds were generated using the following code:

```
DATA seeds;
DO i=1 to 3;
seed=int(10000*ranuni(214366));
OUTPUT;
END;
RUN;
```

If additional seeds are required, then the initial seed as specified will be incremented by 1.

6.3.2. Count Data Model Checking

Distributional assumptions underlying the model used for analysis will be checked. If there are any departures from the distributional assumptions, alternative models may be explored.

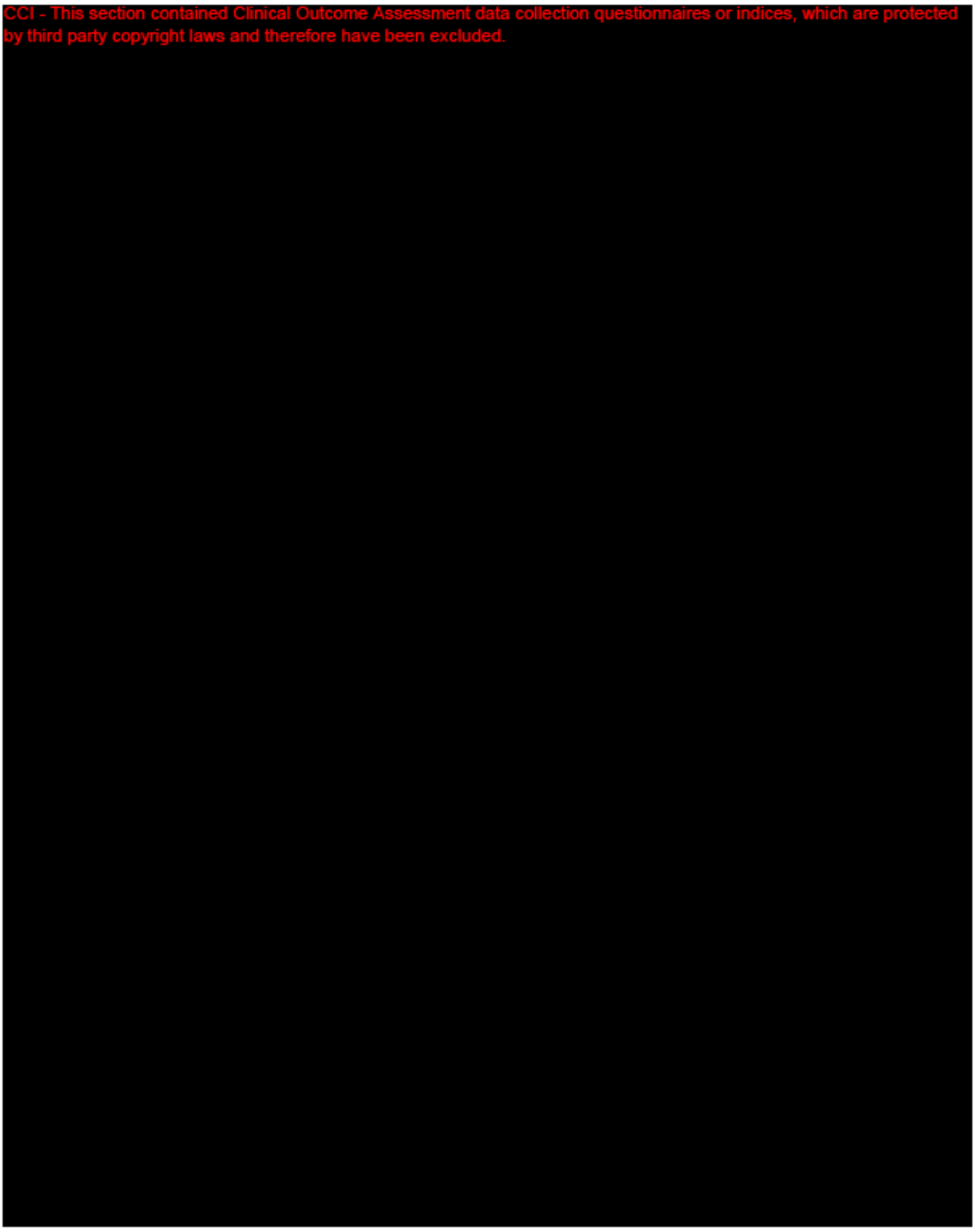
6.3.3. Continuous Data Model Checking

Distributional assumptions underlying the model used for analysis will be examined by obtaining a normal probability plot of the residuals and a plot of the residuals versus the fitted values (*i.e*. checking the normality assumption and constant variance assumption of the model respectively) to gain confidence that the model assumptions are reasonable.

If there are any departures from the distributional assumptions, alternative models will be explored using appropriate transformed data.

6.4. Appendix 4: Flu-PRO Plus responder definition

In September 2020, FDA released Guidance (https://www.fda.gov/media/142143/download) encouraging the development and use by sponsors of a responder definition. To develop this responder definition Flu-PRO Plus items have been mapped onto the item suggested by the FDA in their guidance. The mapping is as follows:



CCI - This section contained Clinical Outcome Assessment data collection questionnaires or indices, which are protected by third party copyright laws and therefore have been excluded.

A responder will be defined as

- Items CCI and CCI scoring CCI
- Items CCI scoring no more than "CCI" in severity
- Items CCI scoring CCI
- Items CCI scoring CCI or CCI

A sustained response is defined as sustained for >= 48hrs.

7. REFERENCES

Brand, J. P. L. Development, Implementation, and Evaluation of Multiple Imputation Strategies for the Statistical Analysis of Incomplete Data Sets. Ph.D. thesis, Erasmus University 1999.

Carpenter, J.R. & Kenward M.G. Multiple Imputation and its Applications. Wiley 2012

Hwang IK, Shih WJ, DeCani JS, Statistics In Medicine. Dec 1990 (Volume9, Issue 12, 1439-1445) Group sequential designs using a family of type I error probability spending functions.

Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika 1983;70:659-63.

Permutt T, Li F. Trimmed means for symptom trials with dropouts. Pharm Stat 2017;16: 20-8.

Sanofi. Press release. https://www.sanofi.com/en/media-room/press-releases/2020/2020-04-27-12-58-00.

Van Buuren, S. Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification. Statistical Methods in Medical Research 2007; 16:219‑242.

Vir Document: VIR-7831-5001 Original Protocol (GSK Study ID 214367). A randomized, multi-center, double-blind, placebo-controlled study to assess the safety and efficacy of monoclonal antibody VIR-7831 for the early treatment of coronavirus disease 2019 (COVID-19) in non-hospitalized patients. Effective Date: 27 July 2020