Official Title: A Phase II, Open-Label, Randomized, Multicenter Study to Investigate The Pharmacodynamics, Pharmacokinetics, Safety, And Efficacy of 8 Mg/Kg or 4 Mg/Kg Intravenous Tocilizumab in Patients With Moderate to Severe Covid-19 Pneumonia

NCT Number: NCT04363736

Document Date: SAP Version 2: 01-September-2020

STATISTICAL ANALYSIS PLAN

TITLE:	**A PHASE II, OPEN-LABEL, RANDOMIZED, MULTICENTER STUDY TO INVESTIGATE THE PHARMACODYNAMICS, PHARMACOKINETICS, SAFETY, AND EFFICACY OF 8 mg/kg OR 4 mg/kg INTRAVENOUS TOCILIZUMAB IN PATIENTS WITH MODERATE TO SEVERE COVID-19 PNEUMONIA**
PROTOCOL NUMBER:	CA42481
STUDY DRUG:	Tocilizumab (RO4877533)
VERSION NUMBER:	2
IND NUMBER:	148225
EUDRACT NUMBER:	2020-001154-22
SPONSOR:	F. Hoffmann-La Roche Ltd.
PLAN PREPARED BY:	[redacted], Ph.D.
DATE FINAL:	Version 1: 19 June 2020
DATE AMENDED	Version 2: See electronic date stamp below

STATISTICAL ANALYSIS PLAN APPROVAL

Date and Time(UTC)	Reason for Signing	Name
01-Sep-2020 12:30:07	Company Signatory	[redacted]

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

The Statistical Analysis Plan version 2 for Study CA42481 was amended from version 1 as follows:

1. Time to event analyses clearly estate that proportional hazard models for both cause-specific hazards and subdistribution hazards will be implemented according to Grambauer et al. (2010)
2. Additional measures of treatment effect were added for the ordinal scale: 1) competing probability over time and 2) median of pair-wise differences
3. Time to event endpoints were changed from “time from randomization” to “time from first dose of study drug”
4. The censoring rules for the time to event endpoints were updated
5. Removed error mentioning placebo-treated group in the analysis of ordinal data
6. Added section specifying organ failure-free days (OFD) endpoint
7. The handling of patients who withdraw not during hospitalization in the analysis of the time to clinical failure was clarified
8. Added statement that for pharmacokinetic (PK), pharamcodynamic (PD) and efficacy, estimates will be provided within strata, and also, when appropriate, across strata, i.e., Competing Probability for the Ordinal Scale.
9. Added an appendix describing handling of below the limit of quantification (BLQ) data in PK and PD
10. Added an appendix with additional time window descriptions
11. Added clarification on what to consider time to clinical improvement

Additional minor changes have been made to improve clarity and consistency.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

GLOSSARY OF ABBREVIATIONS

Abbreviation	Definition
AE	adverse event
AEGT	adverse event grouped term
AESI	adverse events of special interest
BAL	Broncho alveolar lavage
BAP	Biomarker Analysis Plan
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
COVID-19	coronavirus 2019/ SAS-CoV-2
CPAP	Clinical Pharmacology Analysis Plan
CRP	C-reactive protein
ECMO	extracorporeal membrane oxygenation
eCRF	electronic Case Report Form
ICU	intensive care unit
IL-6	interleukin 6
IL-6R	interleukin-6 receptor
IV	intravenous
LDH	lactose dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
NEWS2	National Early Warning Score 2
OFD	Organ failure-free days
PCR	polymerase chain reaction
PD	pharmacodynamic
PK	pharmacokinetic
PT	preferred term
Pt	prothrombin time
RT-PCR	reverse-transcriptase polymerase chain reaction
SAE	serious adverse event
SAP	Statistical Analysis Pan
sIL-6R	soluble interleukin-6 receptor
SMQ	Standard MedDRA Query
SMT	study management team

Abbreviation	Definition
SoC	standard of care
SOC	System Organ Class
SpO2	blood oxygen saturation
TB	tuberculosis
TCZ	Tocilizumab
TTCI	time to clinical improvement
ULN	Upper limit of normal
VFDs	ventilator-free days
WHO	World Health Organization

1. BACKGROUND

This Statistical Analysis Plan (SAP) provides details of the planned analyses and statistical methods for the clinical efficacy and clinical safety for Study CA42481. The primary pharmacodynamic (PD) and pharmacokinetic (PK) objectives will be covered in a separate Clinical Pharmacology Analysis Plan (CPAP). Any analyses of additional biomarkers will be covered in a separate Biomarker Analysis Plan (BAP).

There are currently no drugs licensed for the treatment of patients with severe acute respiratory syndrome corona virus 2 (SARS-CoV-2) (COVID-19). Based on the results from an initial 21-patient retrospective observational study, in which patients with severe COVID-19 pneumonia were treated with tocilizumab (TCZ) off-label (Xu et al. 2020), along with standard-of-care (SoC) treatment, could provide efficacy, offering the potential benefit to treat COVID-19 in hospitalized populations; with the limitation for this observational study of a lack of a proper control as a comparator.

The COVACTA (Study WA42380) study (clinicaltrial.gov number: NCT04320615) was designed to validate the previous findings. COVACTA is a Phase III, randomized, double-blind, placebo-controlled, multicenter study to assess the efficacy and safety of TCZ in combination with SoC compared with matching placebo in combination with SoC in hospitalized adult patients with severe COVID−19 pneumonia. Nevertheless, the optimal dose of TCZ in the treatment of COVID−19 is not known. The COVACTA study is assessing the 8-mg/kg TCZ dose, but Xu et al. (2020) used 400 mg dose (which equates to between 4 and 8 mg/Kg based on body weight range of the Chinese adult population). Study CA42481 will compare two different doses of TCZ: 4 mg/kg and 8 mg/kg.

2. STUDY DESIGN

Study CA42481 is a Phase II, open-label, randomized, multicenter study to assess the pharmacodynamics, pharmacokinetics, safety and efficacy of two different doses of TCZ in combination with SoC in hospitalized adult patients with moderate and severe COVID-19 pneumonia.

The Sponsor intends to enroll approximately 100 patients who have been hospitalized with COVID-19 pneumonia. The study is open-label and is unblinded. Blinding was not considered necessary given that the main objectives of the study involve PD and PK (such as interleukin-6 [IL-6], interleukin-6 receptor [IL-6R], and C-reactive protein [CRP]) outcomes.

Patients must be at least 18 years old with confirmed SARS-CoV-2 (COVID-19*)* infection per the World Health Organization (WHO) criteria, including a positive polymerase chain reaction (PCR) of any specimen (e.g., respiratory, blood, urine, stool, other bodily fluids). At the time of enrollment, patients are categorized as having severe or moderate COVID-19 pneumonia. Severe patients must have blood oxygen saturation (SpO_2)

≤93% or PaO_2/FiO_2[1] <300 mmHg despite being on SoC, which may include anti-viral treatment, low-dose steroids, and supportive care. Moderate patients do not need to meet these oxygen requirements but must have elevated CRP levels to at least 2 times the upper limit of normal (ULN).

Patients in whom, in the opinion of the treating physician, progression to death is imminent and inevitable within the next 24 hours, irrespective of the provision of treatments, will be excluded from the study. Patients with active tuberculosis (TB) or suspected active bacterial, fungal, viral, or other infection (besides COVID-19) will be excluded from the study.

Patients will be randomized at a 1:1 ratio to receive open-label TCZ treatment with either 4 mg/kg or 8 mg/kg in addition to SoC per local practice. Randomization will be stratified by disease status: moderate or severe. The proportion of patients with moderate symptoms will be capped at no more than 50% of the overall study population.

For both arms, if the clinical signs or symptoms worsen or do not improve (e.g., a patient has a sustained fever or experiences clinically significant worsening of signs or symptoms such as an increased supplemental oxygen requirement,), one additional infusion of unblinded treatment of either 4 or 8 mg/kg can be given within 8 to 24 hours after the initial infusion. The maximum dose per infusion is 800 mg.

Patients who do not meet the criteria for participation in this study (screen failure) may qualify for one re-screening opportunity (for a total of two screenings per participant) at the investigator's discretion. The investigator will record the reasons for screen failure in the screening log.

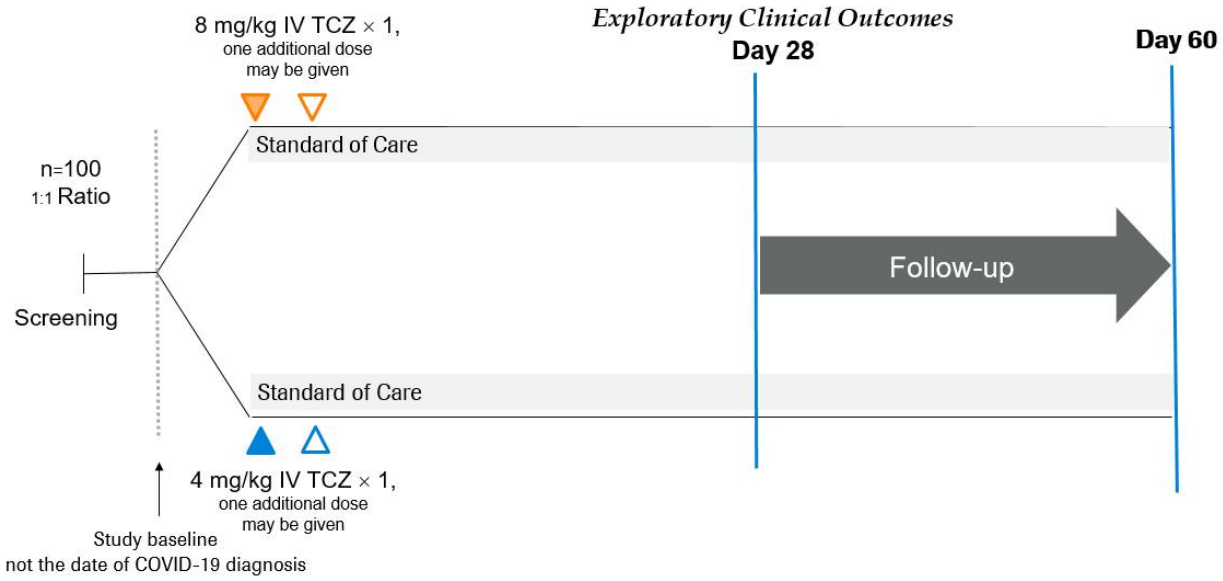
The study assessments to be conducted include the following: physical examination, vital signs, oxygen saturation, assessment of consciousness, presence and absence of respiratory support, adverse events (AEs), concomitant therapies, clinical laboratory tests, and nasopharyngeal swabs. Please see Appendix 2, Appendix 3, and Appendix 4. for details concerning the timing of these assessments.

The Study CA42481 protocol was amended with substantial changes to the outcomes. All outcomes in this SAP are exploratory. Population PK modelling and biomarker analyses will be dealt with in additional plans.

Figure 1 presents an overview of the study design. The schedules of activities are provided in Appendix 2, Appendix 3, and Appendix 4.

[1] The ratio between the blood pressure of the oxygen (partial pressure of oxygen, PaO_2) and the percentage of oxygen supplied (fraction of inspired oxygen, FiO_2).

Figure 1 Study CA42481 Schema



IV = intravenous; TCZ = tocilizumab.

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is provided in Appendix 1. For additional details, see the schedule of activities in Appendix 2, Appendix 3, and Appendix 4.

2.2 OBJECTIVES

This study will evaluate the pharmacodynamics, pharmacokinetics, safety, and efficacy of two doses of TCZ in combination with SoC in hospitalized patients with severe or moderate COVID-19 pneumonia. Specific objectives and corresponding endpoints for the study are outlined below.

For PK, PD and efficacy, results will be given within strata and as pooled estimates across strata when appropriate.

2.2.1 Pharmacodynamic Objective

The PD objective for this study is to assess the differences between two doses of TCZ in combination with SoC for the treatment of moderate to severe COVID-19 pneumonia on the basis of the following endpoints:

- Serum concentrations of IL-6, sIL-6R, ferritin, and CRP following administration of 8 mg/kg and 4 mg/kg intravenous (IV) TCZ at specified timepoints

2.2.2 Pharmacokinetic Objectives

The PK objective for this study is to characterize the pharmacokinetics of two different doses of TCZ in combination with SoC for the treatment of moderate to severe COVID-19 pneumonia on the basis of the following endpoints:

- Serum concentrations of TCZ following administration of 8 mg/kg and 4 mg/kg IV TCZ at specified timepoints (see Appendix 2, Appendix 3, and Appendix 4)

The characterization of serum concentrations at pre-specified timepoints of 4 mg/kg and 8 mg/kg in combination with SoC as well as the association between pharmacokinetics and pharmacodynamics of TCZ over all timepoints will be described in a separate CPAP.

2.2.3 Exploratory Efficacy Objectives

All efficacy objectives in this study are exploratory, hence differences and 95% confidence intervals (CIs) of differences but not p-values will be reported. The objectives will compare the efficacy of 4-mg/kg TCZ dose against 8-mg/kg TCZ dose in combination with SoC for the treatment of severe and moderate COVID-19 pneumonia. The main analysis population will be the modified intent-to-treat (mITT) population defined in Section 4.1.1, and the 4-mg/kg TCZ dose will contain both single and double doses of 4mg/kg presented by randomized group. Likewise, the 8-mg/kg TCZ dose will be treated similarly. Further exploratory efficacy analyses may use the safety population to compare actually received treatments, i.e., 4, 4+4, 8, combined (4+4 and 8), and 8+8.

The National Early Warning Score 2 (NEWS2) values will be calculated by the Sponsor based on vital sign parameters and NEWS2-related assessments recorded by the investigator on the appropriate electronic Case Report Form (eCRF).

In addition to vital measurements, the patient's consciousness level and the presence or absence of respiratory support will be recorded. The NEWS2 parameter for respiratory support is the selection of either "air" or "oxygen," which can include other forms of ventilation to maintain oxygen saturation (see Appendix 5).

The following endpoints are considered:

- Clinical status as assessed using a 7-category ordinal scale over time and specifically at Days 7, 14, and 28

  The ordinal scale categories are as follows:

  1. Discharged (or "ready for discharge" as evidenced by normal body temperature and respiratory rate, and stable oxygen saturation on ambient air or≤2 L supplemental oxygen)
  2. Non-intensive care unit (ICU) hospital ward (or "ready for hospital ward") not requiring supplemental oxygen
  3. Non-ICU hospital ward (or "ready for hospital ward") requiring supplemental oxygen

4. ICU or non-ICU hospital ward, requiring non-invasive ventilation or high-flow oxygen
5. ICU, requiring intubation and mechanical ventilation
6. ICU, requiring extracorporeal membrane oxygenation (ECMO) or mechanical ventilation and additional organ support (e.g. vasopressors, renal-replacement therapy)
7. Death

- Time to clinical improvement (TTCI), defined as a NEWS2 score of ≤ 2 maintained for 24 hours
- Time to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical status. Patients that are in category 2 at baseline will need to achieve category 1 to meet the endpoint of recovery.
- Duration of supplemental oxygen
- Incidence of ICU stay
- Incidence of mechanical ventilation
- Duration of ICU stay
- Time to clinical failure, defined as the time to death, mechanical ventilation, ICU admission, or withdrawal (whichever occurs first). For patients entering the study already in ICU or on mechanical ventilation, clinical failure is defined as a one-category worsening on the ordinal scale, withdrawal or death.
- Mortality at Day 7, 14, 21, 28 and day 60 will be summarized descriptively.
- Time to hospital discharge or "ready for discharge" (as evidenced by normal body temperature and respiratory rate, and stable oxygen saturation on ambient air or ≤ 2 L supplemental oxygen)
- Frequency of the addition of a second dose of TCZ (either 8 mg/kg or 4 mg/kg) at the discretion of the treating physician
- Organ failure-free days (OFD) from baseline to Day 28
- Ventilator-free days (VFD) from baseline to Day 28

Withdrawals are patients for which a discontinuation form is completed. Withdrawals for different reasons, e.g. lack of efficacy vs. administrative censoring, are not distinguished.

2.2.4 Biomarkers

The exploratory biomarker objectives for this study are to further evidence of TCZ pharmacological activity (i.e., PD biomarkers), and overall increase our knowledge and understanding of disease pathogenesis and drug safety, on the basis of the following endpoint:

- Exploratory analysis of individual biomarkers in relation to efficacy and safety (listed in Section 4.5.6, "Laboratory, Biomarker, and Other Biological Samples" of the protocol), and in both blood- and tissue-derived samples.

A separate BAP details how data collected in CA42481 will be used to enhance our understanding of PD, and Efficacy-Response biomarkers of TCZ in the moderate and/or severe COVID-19 pneumonia population.

2.2.5 Safety Endpoints

The safety objective for this study is to compare the safety profile of 4 mg/kg TCZ dose against that of 8 mg/kg TCZ dose in combination with SoC for the treatment of severe and moderate COVID-19 pneumonia on the basis of the following endpoints:

- Incidence and severity of AEs, with severity determined according to National Cancer Institute Common Terminology Criteria for AEs (NCI CTCAE) v5.0
- SARS-CoV-2 (virus causing COVID-19) viral load over time, as collected by nasopharyngeal swab and bronchoalveolar lavage (BAL) samples (if applicable)
- Time to reverse-transcriptase polymerase chain reaction (RT-PCR) virus negativity
- The proportion of patients with any post-treatment infection
- Change from baseline in targeted clinical laboratory test results

2.3 DETERMINATION OF SAMPLE SIZE

The purpose of this open-label study is the comparison of TCZ pharmacokinetics and PK-PD relationship between two doses of TCZ (8 and 4 mg/kg) in hospitalized patients with moderate to severe COVID-19 pneumonia. The comparison will be assessed using modeling and evaluating the safety and efficacy measures between Day 1 and Day 28 post-randomization.

Data from the follow-up period from Days 28 to 60 may also be assessed. Approximately 100 patients will be enrolled and randomized in a 1:1 ratio to the 8-mg/kg and 4-mg/kg IV TCZ arms in this study. Disease severity type will also be stratified into moderate and severe COVID pneumonia, giving an overall equal four-way split between the two doses and type of disease severity and resulting in four groups of approximately 25 patients each. Patients may receive a repeat dose within 8 to 24 hours after the first TCZ dose if clinicians judge it is necessary (resulting in a TCZ dose range from 4 to 16 mg/kg, with a maximum dose of 800 mg). The variability of PK parameters of TCZ in a heterogeneous COVID-19 population in an acute setting is unknown; however, 20 to 30 patients is deemed adequate considering the between-subject variability in healthy volunteers and patients with rheumatoid arthritis (RA) provided by population-PK modeling. Data quality issues, patient dropouts, and early hospital discharge may all lead to missing data, especially in a pandemic setting; thus, enrolling 50 patients according to dose and disease severity type (25 per arm) may offset missing data and allow an adequate volume of PK and PD data to inform the population-PK model and allow a comparison of the 4- and 8-mg/kg TCZ doses.

2.4 ANALYSIS TIMING

The primary study analysis will occur when the last patient enrolled has completed study visit Day 28 or has withdrawn, and will be based on cleaned data for all patients up to and including this point.

There will be an additional analysis on the final data when all patients have either completed study visit Day 60 or withdrawn and all data are available. Analyses from the first reporting event, restricted to data up to Day 28, will not be updated based on the final snapshot.

3. STUDY CONDUCT

This study will enroll approximately 100 patients that have been diagnosed with COVID-19 pneumonia and meet the entry criteria in United States centers. Patients will be randomized at a 1:1 ratio to receive unblinded treatment with either 4 or 8 mg/kg TCZ in addition to SoC. For both arms, if the clinical signs or symptoms worsen or do not improve, one additional infusion of the original TCZ concentration can be given within 8 to 24 hours after the initial infusion.

Patients who do not meet the criteria for participation in this study (screen failure) may qualify for one re-screening opportunity (for a total of two screenings per participant) at the investigator's discretion.

Patients will be followed up for a total of 60 days after first dose of study medication.

If patients are discharged from hospital prior to Day 28, follow-up visits should occur twice weekly through Day 28. Patients are encouraged to return for onsite visits at least weekly if possible; other follow-up visits may be conducted by telephone or home visits. Patients should return to the site for a Day 28 visit. After Day 28, all patients should have follow-up visits on Day 35, Day 45, and Day 60; the Day 35 and Day 45 visits may be conducted by telephone or by home visits for discharged patients, while the Day 60 visit should be conducted onsite.

During the study, standard supportive care will be given according to clinical practice.

3.1 RANDOMIZATION, STRATIFICATION AND BLINDING

Patients will be randomized as soon as possible after screening at a 1:1 ratio to receive unblinded treatment of either 4 or 8 mg/Kg TCZ, respectively. Study treatment must be given in combination with SoC. The randomization will be stratified by disease severity (severe or moderate); and will occur through use of a permuted-block randomization method to ensure a balanced assignment to each treatment arm. The proportion of patients with moderate symptoms at the time of randomization will be capped at no more than 50% of the overall study population. The study is open-label and is unblinded; thus

the study team will monitor safety on an ongoing basis, i.e., there will be no independent Data Monitoring Committee.

3.2 DATA MONITORING

The study management team (SMT) will monitor the incidence of all serious adverse events (SAEs), adverse events of special interest (AESI) and any anticipated events during the study.

Regular safety reviews will begin after approximately 10 to 20 patients have been enrolled and reached Day 7 of the study. Following this and additional safety reviews, the SMT may recommend to continue the study without modifications, modify the conduct of the study, or discontinue the study as described in Section 4.6.3 Study Discontinuation, of the protocol.

4. STATISTICAL METHODS

4.1 ANALYSIS POPULATIONS

Disposition summaries will be based on an all-patient population (all patients randomized and/or receiving study drug). Efficacy analyses will be based on the mITT population, if not otherwise specified. Analysis of safety data, and PD data will be based on the safety population.

4.1.1 Modified Intent-to-Treat Population

The mITT population is defined as all patients randomized in the study who receive any amount of study drug, with patients grouped according to the treatment assignment at randomization.

4.1.2 Safety Population

Safety population will consist of all patients who receive any amount of study drug. In all safety, PK and PD analyses, patients will be grouped according to the first treatment that patients receive rather than the treatment assigned at randomization.

4.1.3 Pharmacokinetics Population

The PK population is defined as the subset of the safety population with at least one valid PK result.

4.2 ANALYSIS OF STUDY CONDUCT

The number of patients enrolled, discontinued, or who complete the study will be summarized to Day 28 and to the end of the study. Reasons for premature study discontinuation will be listed and summarized to Day 28, as well as the end of the study. Listing of randomized patients and a listing of investigators will be produced.

The number of patients discharged from hospital will also be summarized by visit.

Eligibility criteria and other major protocol deviations will be listed and summarized by treatment group.

The patients excluded from the safety and mITT populations will be summarized, including the reason for exclusion by treatment group. A summary of enrollment by investigator name will be produced.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Demographic and baseline characteristics (including, but not limited to, age, sex, race, geographic region, NEWS2, ordinal scale for clinical status, IL-6, sIL-6R, mechanical ventilation, anti-viral treatment at baseline, steroids at baseline) will be summarized using means, standard deviations, medians, and ranges for continuous variables and proportions for categorical variables, as appropriate. Summaries will be presented by treatment group and will be presented for the mITT population and may, in addition, be presented for the safety population.

4.3.1 Demographics

The key demographic parameters are the following:

- Sex
- Age
- Weight
- Race
- Ethnicity
- Female fertility status
- Smoking history (never, current, or former)
  - Former/current user: number of years a patient smoked (years), nicotine exposure reported in pack years
  - E-cigarettes use (yes or no)

4.3.2 Disease Characteristics

The key disease characteristics are:

- Disease severity
- NEWS2
- Ordinal scale for clinical status
- IL-6
- sIL-6R
- Ferritin
- CRP
- D-dimer

- Prothrombin Time (Pt)
- activated Partial Thromboplastin Time (aPTT)
- Platelets
- lactose dehyodrogenase (LDH)
- WBC (including neutrophils % and count –ANC-)
- Lymphocyte count and %.
- Median time from symptom onset to first dose
- Median oxygen saturation or PaO2:FiO2 ratio
- Mechanical ventilation (levels 5-6 of ordinal scale for clinical status)
- Admission to ICU

  Use of continuous renal-replacement therapy
- Steroid use for up to 7 days prior to randomisation (to be derived from concomitant medication)
- Anti-viral treatment for up to 7 days prior to randomisation (to be derived from concomitant medication)
- Symptoms at time of COVID-19 diagnosis
  - Fever
  - Cough
  - Shortness of breath
  - Gastrointestinal symptoms (e.g., diarrhea, nausea and loss of appetite)
  - Headache
  - Fatigue
  - Other
- Number of days from first COVID-19 symptom at baseline
- COVID-19 diagnosis based on PCR of specimen type
- Number of days from first COVID-19 diagnosis at baseline
- Specimen type at screening
- Quantitative PCR result
- Overall SARS-CoV-2 antibodies

Central laboratory results will be used whenever available. Only in the case of missing central laboratory results, local laboratory results could be used.

4.3.3 Medical History

Medical history data will be summarized descriptively by treatment group using the safety population. A glossary showing the mapping of investigator verbatim terms to diseases will be produced for the medical history data.

4.3.4 Surgeries and Procedures

A listing of any previous or ongoing surgeries and procedures will be produced for the safety population.

4.3.5 Previous and Concomitant Medications

Previous and concomitant treatments will be summarized descriptively by treatment group for the safety population. Previous treatments that have been stopped prior to study Day 1 will be summarized separately. There will be a summary of all concomitant treatments, including those that were initiated prior to study Day 1. In addition, there will be a summary of all anti-viral treatments, as previous medical treatment and/or concomitant medication, with the indication given as COVID-19, COVID prophylaxis, COVID pneumonia, pneumonia or similar. Any of these treatments categorized under the Medication Class of Investigational Drug will be summarized with the proximity (up to 7 days prior randomization) of administration relative to tocilizumab.

A glossary showing the mapping of investigator verbatim terms to medication-coded terms will be produced for previous or concomitant medication.

4.4 VISIT LABELS

For summaries of data not collected by visit, such as AEs, medical history, and concomitant medications all data up to the end of study will be included. Exceptions to this include death and discharge; which will be summarized weekly in descriptive summaries, following the time windowing approach described below.

Deaths will also be captured on the ordinal scale of clinical status, where the completion of an unscheduled ordinal scale value of 7 will be recorded in eCRF. Deaths confirmed by public record are also captured on the eCRF, which may not have been captured as AEs for patients who are withdrawn from the study. These events will also be incorporated into the windowing for death.

Patient assessments that are collected at scheduled visits will be assigned to a study visit using the actual study day of the assessment; this includes data from withdrawal visits and any unscheduled visits. Time windows will be continuous from the midpoint between two consecutive study visits to the next midpoint, and will be dependent on the schedule of assessments for each variable independently. An example of time windowing for the PD biomarkers is shown below (Table 1).

Table 1 Time Windows for Assigning Assessment Study Days to Study Visits for PD Parameters

Scheduled Study Day	Efficacy Time Window (days) [a]
Day 1 (Baseline)[b]	≤1
Day 2	2
Day 3	3
7 (Week 1)	>3 to ≤10
14 (Week 2)	>10 to ≤17
21 (Week 3)	>17 to ≤24
28 (Week 4)	>24 to ≤28
35 (Week 5)	>28 to ≤38
Day 60	>38 to ≤67

[a] From Week 1 onward use value nearest to scheduled study day. Day 1 includes at least two samples (one at baseline and the second 15 minutes after infusion).

[b] A second postdose sample on Day 1 will be taken from patients given a second dose of TCZ (see SOA).

Where there is more than one efficacy assessment within a time window, then the non-missing assessment nearest in time to the expected time window will be assigned to that visit. If two or more assessments are equidistant from the scheduled timepoint, then the latest assessment will be used for efficacy (other than death, in which case the assessment prior to the visit week will be used).

For safety parameters such as laboratory parameters and vital signs the “worst case” will be used.

The last value from screening will be used for baseline assessments if there is no baseline (study Day 1) value. Pretreatment assessments will be used preferentially on study Day 1 for baseline.

4.5 PHARMACODYNAMIC, PHARMACOKINETIC, AND EFFICACY OBJECTIVES

4.5.1 Pharmacodynamic and Pharmacokinetic Objectives

The main PD endpoints in this study are IL-6, sIL-6R, ferritin and CRP levels measured at baseline and then at Days 1, 3, 7, 14, 21, and 28 post-randomization. The median, standard error of the median, minimum and maximum values will be summarized by day for each dose using the mITT population. The median of all pairwise differences between the 4 mg/kg and the 8 mg/kg will also be reported for the mITT population. Standard error of medians will be obtained via bootstrapping. Results will be reported with both tables and plots.

The population is the safety population defined in this study. The estimate is the median difference in IL6, sIL-6R, CRP and ferritin levels between treatments across all visits from baseline to Day 28.

Likewise, PK data will be summarized similarly over time using the PK population.

This information will help inform prescribers about the potentially similar pharmacodynamics effects of a 4-mg/kg dose compared to an 8-mg/kg dose of TCZ in COVID-19 patients.

4.5.2 Efficacy Endpoints

All efficacy endpoints are exploratory. All efficacy analyses will use the mITT population.

Sensitivity analyses to evaluate the robustness of results to the analysis methods (e.g., handling of withdrawals) may be conducted and are described in this SAP in each relevant section.

Descriptive subgroup analyses to evaluate the consistency of results across prespecified subgroups may also be conducted as specified in Section 4.5.4.

4.5.2.1 Ordinal 7-Category Scale

The median clinical status according to the 7-category ordinal scale and its 95% CI via bootstrapping for the 4-mg/kg and the 8-mg/kg will be tabulated and plotted for all days between baseline and Day 28. The magnitude of the difference between doses will also be assessed at Days 7, 14 and 28 via the van Elteren test adjusted for baseline stratification factors and by a proportional odds model where the odds ratio and 95% CI will be reported. Also the median of all pair-wise differences (Hodges-Lehmann estimator) between doses will be plotted over time with 95% CI obtained by bootstrapping. All analyses will use the mITT population.

The assumption of proportional odds will be evaluated by visually comparing the fitted proportions of patients across the ordinal scale from the model with the observed data.

The competing probability of observing patients on the 4 mg/kg dose having ordinal scores higher than patients on the 8 mg/kg dose plus half of the probability of having equal scores will be plotted daily with 90% CIs (Zhao et al. 2012).

In addition to imputing the ordinal scale at Day 28 with an earlier death or discharge (without re-admittance), this imputation rule will also be followed at earlier timepoints. A death or discharge (unless the patient is re-admitted) will always be carried forward to all subsequent assessments regardless of what is recorded for the ordinal scale. If discharge is to a nursing facility then the ordinal scale should be still collected, otherwise it will be set to missing. For patients who withdraw before Day 28, their last post baseline ordinal category prior to withdrawal will be used in the primary analysis, unless death within the time frame was captured from public records or otherwise; in which case death will be used in the analysis.

Moreover, the ordinal scale will be summarized by visit and treatment showing n and percentage in each category, as well as missing data. A stacked bar chart of the ordinal scale will be produced by treatment group, the bars will total to 100% and the categories, including “missing,” will be shown.

Assessment of patient status using this ordinal scale will be recorded at baseline and once daily in the morning (between 8 a.m. and 12 p.m.) while hospitalized.

Additionally, for patients that withdraw prior to Day 28, the 50th percentile of the data from those that complete the study to Day 28 will be used as imputed value.

4.5.2.2 Time-to-Event Analyses

Time-to-event endpoints will be compared between the low and high TCZ doses using the stratified (moderate vs. severe) proportional hazard models using the mITT population. In all analyses apart from time to clinical failure, death will be a competing risk. Models for the cause-specific hazards, i.e., Cox model, and also for the subdistribution hazards (Fine and Grey 1999) will be used. The subdistribution hazards model assumes death to be right censored and death times happening after 28 days on trial (672 hours) minus the number of hours from midnight to the time of first treatment in each patient. Non-parametric cumulative incidence functions for the event of interest, with 95% CIs, will be obtained using the Aalen-Johansen estimator. In addition to hazard ratios, the difference in cumulative incidence curves between the low and high TCZ doses at days 7, 14 and 28 will be provided.

For time to event endpoints that include discharge as an event, the earliest time of discharge or “ready for discharge” from the different sources of discharge will be used in all analyses. If a patient is discharged and re-admitted more than 12 hours later, then the first discharge will be considered as meeting the event. If a patient is readmitted within 12 hours of discharge, then they will not have met the endpoint at this time. If

they are discharged later in the study (without re-admittance within 12 hours), then the later time of discharge will be used.

Time-to-event endpoints include:

- TTCI in hours

  Defined as time from first dose of study drug to NEWS2 score of ≤ 2 maintained for 24 hours

Intercurrent events are those that occur after treatment initiation and either preclude observation of the variable or affect its interpretation. Intercurrent events, such as patients who are lost to follow-up, die or discontinue for any reason prior to achieving the event or do not have the event, will be accounted for through censoring rules, as described in Table 2.

Partial date times may be imputed based on available data, following a conservative approach. The NEWS2 is to be assessed twice daily, with approximately 12 hours between each assessment. At least two scheduled assessments with a score of ≤ 2 covering a span of at least 21.5 hours, with a maximum of 26.5 hours between the first and last of these assessments (there must be no assessments with a score >2 in between), will be required to meet the criterion. If a patient has a score of ≤ 2 and is then discharged from hospital within 26.5 hours, with no subsequent scores >2 before the discharge they will have met the endpoint.

Table 2 Time to Clinical Improvement and Censoring

Event	Censor	Date and Time
Hospital discharge prior to clinical improvement criterion met	Yes	Hospital discharge
Hospital discharge and hospital re-admission within 12 hours and continue study	No	Not applicable
Death prior to clinical improvement criterion met	No (competing risk)	Death
Discontinuation or lost to follow-up for any reason prior to clinical improvement criterion met	Yes	Last scheduled vital sign assessment
No clinical improvement	Yes	Last vital sign assessment within Day 28 time window

Other time-to-event endpoints include the following:

- Time to improvement in ordinal clinical status (days)

  Defined as time from first dose of study drug to the time when at least a 2-category improvement in the 7-category ordinal scale is observed.

For patients in category 2 at baseline, discharge will be considered as meeting the threshold. For patients who are discharged and the ordinal scale assessment has not been completed at discharge, they will be assumed to be in category one of the ordinal scale at the point of discharge, unless they are re-admitted within 12 hours. Intercurrent events, such as patients who are lost to follow-up, die or discontinue for any reason prior to achieving the event or do not have the event, will be accounted for through censoring rules, as described in Table 3.

Table 3 Time to Improvement in Ordinal Clinical Status and Censoring

Event	Censor	Date and Time
Hospital discharge and hospital re-admission within 12 hours and continue study	No	Not applicable
Death prior to improvement in ordinal clinical status criterion met	No (competing risk)	Death
Discontinuation or lost to follow-up for any reason prior to improvement in ordinal clinical status criterion met	Yes	Last ordinal scale assessment
No improvement in ordinal clinical status	Yes	Last ordinal scale assessment within Day 28 time window

- Time from first dose of study drug to hospital discharge or “ready for discharge” (hours)

  Ready for discharge; defined as normal body temperature and respiratory rate, and stable oxygen saturation on ambient air or ≤ 2L supplemental oxygen (ordinal scale category one).

Intercurrent events, such as patients who are lost to follow-up, die or discontinue for any reason prior to achieving the event or do not have the event, will be accounted for through censoring rules, as described in Table 4.

Table 4 Time to Hospital Discharge or "Ready for Discharge" and Censoring

Event	Censor	Date and Time
Hospital discharge and hospital re-admission within 12 hours and continue study	No	Not Applicable
Death prior to discharge	No (Competing Risk)	Death
Discontinuation or lost to follow-up for any reason prior to discharge or "ready for discharge" criterion met	Yes	Last ordinal scale assessment
Not discharged or " ready for discharge"	Yes	last Ordinal scale assessment within Day 28 time window

- Time to clinical failure (days)

  Defined as the time from first dose of study drug to first occurrence on study of death, mechanical ventilation, ICU admission or withdrawal (discontinuation from study due to lack of drug efficacy), whichever occurs first. For patients entering the study already in ICU or on mechanical ventilation, clinical failure is defined as a one-category worsening on the ordinal scale, withdrawal or death.

Intercurrent events, such as patients who are lost to follow-up or discontinue for any reason prior to the event or do not have the event, will be accounted for through censoring rules, as described in Table 5 below.

Table 5 Time to Clinical Failure and Censoring

Event	Censor	Date and Time
Hospital discharge not followed by death or re-admittance	Yes	Last scheduled vital sign assessment (or discharge if no post-discharge vital sign data is available)
Hospital discharge and hospital re-admission within 12 hours and continue study	No	Not Applicable
Lost to follow-up prior to clinical criterion met not followed by death	Yes	Last scheduled vital sign assessment or ordinal scale assessment
Clinical failure criterion not met	Yes	Last of scheduled vital sign assessments or ordinal scale assessments within Day 28 time window

The NEWS2 score and clinical failure status as defined above will be summarized descriptively by visit.

4.5.2.3 Incidence Endpoints

Exploratory efficacy incidence endpoints will be analyzed using the Cochran-Mantel-Haenszel (CMH) test statistic adjusted by the stratification factors at baseline disease status severe or moderate using the mITT population, unless stated otherwise. The weighted difference in proportions for the treatment group comparison will be presented, together with a 95% CI (see Appendix 6). All the incidence endpoints will use the same method.

- Incidence of mechanical ventilation by Day 28
- Incidence of ICU stay by Day 28

For incidence of mechanical ventilation or incidence of intensive care stay by Day 28 for patients that have withdrawn or died prior to Day 28, the non-responder rule will be applied, i.e., it will be assumed that the patient required mechanical ventilation, or has had an ICU stay by Day 28 in the analysis. Patients without either mechanical ventilation or intensive care stay, respectively, prior to discharge, will be assumed to be responders in the analysis, unless the patient is re-admitted to the hospital within 12 hours or the patient dies by Day 28.

The number and proportion of patients requiring mechanical ventilation or an ICU stay will be summarized descriptively by study week.

- Difference in mortality at Day 7, 14, 21, 28 and 60

  The difference in proportion of patients that have died by Day 28 will be compared using the CMH test as described above as the main analysis method. All deaths reported post-discontinuation and discharge will be included in these analysis.

  Deaths occurring between each visit, and cumulative deaths by visit will be summarized descriptively to Day 60.

- Frequency of the addition of a second dose of TCZ (either 8 mg/kg or 4 mg/kg) at the discretion of the treating physician

Treating physicians can prescribe a second identical TCZ dose within 8 to 24 hours after the first one if patients do not show signs of improvement. No patient can receive more than a total of 800 mg of TCZ per infusion.

4.5.2.4 Duration Endpoints Ventilator-Free Days

The number of ventilator-free days (VFDs) is defined as the number of days from Day 1 to Day 28 when the patient is alive and breathes without invasive assistance of the mechanical ventilator. VFDs will be derived from the vital signs and oxygen saturation log; if invasive mechanical ventilation or ECMO is recorded for any part of the day, the day will not be counted as a VFD.

VFDs will be zero if the patient is mechanically ventilated from Day 1 to Day 28. VFDs will be zero if a patient dies on or prior to Day 28.

For patients withdrawn early from the study but not discharged, if patients were on invasive mechanical ventilation at the point of discontinuation it will be assumed that the remaining days to Day 28 are not VFDs. For patients not using invasive mechanical ventilation at point of withdrawal, it will be assumed the period to Day 28 are VFDs. For patients that are discharged, days from discharge to Day 28 will be counted as VFDs.

VFDs will be analyzed using the Van Elteren test, including the stratification factors at randomization (disease status: severe vs. moderate). The median VFDs for each treatment group and the corresponding 95% CI for the median will be presented, as well as the difference in medians and a 95% CI for the difference. A cumulative distribution plot of VFDs will be produced.

VFDs during hospitalization will also be summarized descriptively using the medians, with 95% CIs by treatment group for those patients alive at Day 28, with a count of the number of patients assigned zero VFDs due to death by Day 28.

4.5.2.5 Duration of Supplemental Oxygen (Days)

Duration of supplemental oxygen (days) will also be derived from the vital signs and oxygen saturation log, where study days with “supplemental oxygen or other forms of ventilation” will be summed up to and including Day 28. Patients without any supplemental oxygen use will assigned a duration of zero days. For patients withdrawn early from the study but not discharged, if the patient was receiving supplemental oxygen at the point of withdrawal, it will be assumed that the remaining days to Day 28 were on supplemental oxygen. For patients not using supplemental oxygen at point of withdrawal, it will be assumed supplemental oxygen is not required to Day 28. For patients that are discharged, days from discharge to Day 28 will be counted as days without supplemental oxygen (unless supplemental oxygen use is recorded during follow up visits). Duration of supplemental oxygen use will be 28 days if a patient dies on or prior to Day 28.

Days of supplemental oxygen use will be analyzed and summarized descriptively in a similar method to VFDs. In addition, the number and the proportion of patients receiving supplemental oxygen using the observed data will be summarized by visit to Day 60.

4.5.2.6 Duration of ICU Stay (Hours)

Duration of ICU stay (hours) will be calculated as the sum of the number of hours spent in an ICU, up to and including Day 28, based on the admission and discharge date times from the ICU stay information log (ICU discharge date time – ICU admission date time). Multiple periods of ICU stay will be summed. Patients without any ICU stays will be assigned a duration of zero hours.

Partial admission and discharge times may be imputed based on available data, following a conservative approach. For patients that discharged, any ongoing ICU stays without an end date will be imputed from the date of discharge as appropriate and it will be assumed that days from discharge to Day 28 do not involve an ICU stay. For patients not in the ICU at the point of withdrawal from study it will be assumed that the period to Day 28 has no incidences of ICU stay post withdrawal. For patients in ICU on the day of withdrawal it will be assumed that they are in the ICU throughout the period to Day 28. For patients that die on or prior to Day 28 all days will count as an ICU stay.

Hours of ICU stay will be analyzed and summarized descriptively in a similar method to VFDs.

4.5.2.7 Organ failure-free days

Days without organ failure will be summarized descriptively through Week 4. In addition, a summary of individual organ failure over time will be provided.

Organ failure is defined as present on any day when the most abnormal vital signs/abnormal lab value meets the definition of clinically significant organ failure (Bernard et al, 1995; NHLBI ARDS Clinical Trials Network 2014). Cardiovascular and central nervous system function will be assessed through blood pressure and pH (and responsiveness to fluids), and alert, respectively. Pulmonary organ failure will be counted as anyone with PaO2:FiO2 ratio $\leq$ 300 or SpO2 $\leq$ 93% or using supplemental oxygen or on mechanical ventilation. Renal, hepatic and coagulation parameters will be assessed via blood tests in order that the presence of clinically significant organ failure can be determined. Each day a patient is alive and free of a given clinically significant organ failure will be scored as a failure-free day for that organ. In the case of missing data for a particular organ the last observation post-baseline will be carried forward until the next observation. Any day that a patient is alive and free of all 5 organ failures (pulmonary, cardiovascular, renal, hepatic, central nervous system) will be considered OFD.

If a patient dies on or before Day 28, they will be assigned a value of zero OFD. For patients that are discharged, days from discharge to Day 28 will be counted as OFD, unless they are readmitted in which case the available data will be used.

4.5.3 Controlling for Type I Error

All efficacy analyses in this study are exploratory and type I error correction will not be implemented.

4.5.4 Subgroup Analyses

The odds ratio for mortality for 4-mg/kg versus 8-mg/kg doses at Day 28 will be analyzed by logistic regression, including covariates of interest as well as the stratification factors in the model. The odds ratio for the treatment effect by sex, race, age (18–64 years, 65–84 years and $\geq$ 85 years), region and mechanical ventilation will be determined.

Other subgroup analyses may also be performed such as baseline ordinal status 2, 3-4, and 4–5. There should not be patients with baseline score of 6 as those on extracorporeal membrane oxygenation (ECMO), Multi-organ Failure (MOF) and in shock will be excluded.

4.6 SAFETY ANALYSES

Safety assessments will be performed on the safety population. In all safety analyses, patients will be grouped according to the first treatment that the patients received rather than the treatment assigned at randomization.

Descriptive summaries of laboratory values and change from baseline throughout the study will be tabulated by treatment arm. For selected parameters, changes from the proportion of patients experiencing clinically significant changes relative to baseline will be summarized by treatment arm.

Values, along with change from baseline, will be summarized using descriptive statistics for each vital sign parameter.

COVID-19 (SARS-CoV-2) viral load over time, as collected by nasopharyngeal swab and BAL samples (if applicable) will be summarized descriptively by timepoint and treatment group. The number and proportion of patients negative and positive will be displayed, and for those positive the quantitative result will be summarized.

Time to RT-PCR COVID-19 virus negativity will be analyzed using similar methods to those for the other time-to-event analyses.

Overall SARS-CoV-2 antibody concentration will be measured at baseline and Day 28. The median percentage change from baseline will be summarized across groups.

4.6.1 Exposure to Study Drug

Exposure to study drug will be summarized including number of patients with one or two doses and number of patients with dose modification by treatment group.

A listing of patients by treatment group will be prepared detailing dosing of study drug, volume administered and any dose modification.

4.6.2 Adverse Events

Medical Dictionary for Regulatory Activities (MedDRA) will be used as the thesaurus for AEs and disease codes, and the WHO Drug Global B3 Format dictionary will be used for treatments. A glossary of these codes will be produced.

Only treatment-emergent AEs will be summarized. Treatment-emergent events are defined as those AEs with observed or imputed onset date on or after the start date of study treatment. Only where the most extreme severity is greater than the initial severity

(or if most extreme severity is not missing and initial severity is missing) will events with an onset date prior to the start of trial treatment be considered treatment emergent. An AE with a completely missing start date will be assumed to be treatment emergent unless the AE has a complete non-imputed end date that is prior to Day 1.

Adverse events will be coded and tabulated by system organ class (SOC), and/or preferred term (PT) and treatment arm. In tabulations, PTs and their associated SOC will be presented in order of descending frequency summed across the treatment groups.

Adverse events will also be tabulated by severity, as graded according to NCI CTCAE v5.0 scale, and relationship to study medication as indicated by the investigator.

The following will also be summarized:

- SAEs
- AEs leading to withdrawal of study drug
- AEs leading to discontinuation from the study
- AEs leading to death
- Hypersensitivity AEs (AEs occurring immediately after or within 24 hours of the end of an infusion that are not deemed “unrelated” to study treatment)

AESI will be defined using SOC, published Standard MedDRA Queries (SMQs) or Adverse Event Grouped Terms (AEGTs) defined by Roche Drug Safety. The groupings of AEs will include, but may not be limited to, the following:

- Infections (Infections and Infestations SOC)
- Opportunistic infections (Roche Standard AEGT Basket)
- Malignancies (Malignant or Unspecified tumors SMQ Narrow)
- Hepatic events (Hepatic failure, Fibrosis, and Cirrhosis and Other Liver Damage-related Conditions SMQ Wide or Hepatitis, non-infectious SMQ Wide)
- Stroke (Ischemic Cerebrovascular Conditions SMQ Wide or Hemorrhagic Cerebrovascular SMQ Wide)
- Myocardial infarction (MI) (MI SMQ Wide)
- Hypersensitivity reactions (Hypersensitivity SMQ Narrow)
- Anaphylactic reaction events (utilizing Roche Standard AEGT Basket according to Sampson’s criteria) (Sampson et al. 2006) occurring immediately after or within 24 hours of injection of tocilizumab; and a separate summary using the Anaphylactic Reaction SMQ Narrow for events occurring immediately after or within 24 hours of injection of tocilizumab
- Gastrointestinal perforations (Gastrointestinal perforation SMQ Wide)
- Bleeding events (Hemorrhages SMQ Wide)

- Demyelinating events (Demyelination SMQ Narrow)

A glossary showing the mapping of investigator verbatim terms to PTs will be produced for all AEs included in the analysis. For each AE of special interest table based on SMQs/AEGTs, a corresponding listing of the PTs that comprise the SMQ will be produced.

Listings of AEs and SAEs will be produced. AESI will also be listed.

AEs and SAEs will be summarized by age category (18–64, 65–84, and ≥85 years).

The exposure duration on study (exposure duration is the date of the last safety assessment or death if present, minus the date of the first dose of TCZ plus one divided by 365.25) will be summarized.

4.6.3 Laboratory Data

Laboratory values will be converted to Système International units, and data will be transformed to a common Roche standard reference range.

Summary tables will detail the actual values and changes from baseline of the laboratory parameters over visits by treatment arm. Arterial blood gases will be summarized separately. Summaries of the number of patients by NCI CTCAE grade for hematology and hepatic laboratory parameters (alkaline phosphatase [ALP], alanine aminotransferase [ALT], and aspartate aminotransferase [AST]), and total bilirubin) will be produced (for summaries referring to NCI CTCAE grading).

For neutrophils, platelets, lymphocytes and hepatic lab parameters, the number of patients will be summarized by CTCAE grade category for baseline and worst post baseline result.

Patients with values outside the reference will be listed with an indication of the direction of the abnormality (High or Low).

A listing of all pregnancies will be presented.

4.6.4 Vital Signs

Summary statistics on absolute values and their change from baseline for all observed vital signs (diastolic blood pressures, systolic blood pressures, respiratory rate, pulse rate, body temperature, and peripheral oxygen saturation) will be presented over time by treatment group. Baseline is defined as the last assessment prior to treatment. Additionally, a graphical representation of means over time of oxygen saturation and temperature (Daily to Week 4) will be presented.

For patients requiring supplemental oxygen, summary statistics on absolute values of the oxygen flow rate (in liters per minute [L/min]) and/or fraction of inspired oxygen (FiO_2) will be produced by visit, timepoint, and treatment group.

The level of consciousness will be summarized over time.

The number and proportion of patients requiring oxygen supplementation or other form of ventilation will be summarized over time, including type of support given. Non-invasive mechanical ventilation will be summarized overall as well as by its component types (continuous-positive airway pressure [C-PAP], bilevel-positive airway pressure [BiPAP], and other). Invasive mechanical ventilation will also be summarized overall and by component types (endotracheal tube, tracheostomy tube).

A listing of patients with chest X-ray, CT scan, location of opacities, and ECG results (as a separate listing) with clinically significant abnormalities will be produced.

4.7 INTERIM ANALYSES

There are no interim analyses planned in this study.

5. REFERENCES

Bernard GR, Doig G, Hudson LD, et al. Quantification of organ failure for clinical trials and clinical practice. Am J Respir Crit Care Med 1995;151:Suppl:A323-A323.

Fine JP and Grey RJ. A proportional hazards model for the subdistribution of a competing risk. Journal of the American Statistical Association, 1999; 94:496-509.

National Heart, Lung, and Blood Institute ARDS Clinical Trials Network, Truwit JD, Bernard GR, et al. Rosuvastatin for sepsis-associated acute respiratory distress syndrome. N Engl J Med. 2014;370(23):2191-2200.

Sampson HA, Muñoz-Furlong A, Campbell RL, et al. Second symposium on the definition and management of anaphylaxis: summary report—Second National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network Symposium. J Allergy Clin Immunol 2006;117:391–7.

Xu X, Han M, Li, T, et al. Effective treatment of severe COVID-19 patients with tocilizumab. Submitted manuscript. [Resource on the internet]. 2020 [updated: 5 March 2020; cited: 17 March 2020]. Available at: http://www.chinaxiv.org/abs/202003.00026.

Zhao YD, Rahardja D, Dmitrienko A. Estimation of the competing probability for the van Elteren test. Model Assisted Statistics and Applications, 2012; 7:41-47.

Appendix 1 Protocol Synopsis

TITLE:	**A PHASE II, OPEN-LABEL, RANDOMIZED, MULTICENTER STUDY TO INVESTIGATE THE PHARMACODYNAMICS, PHARMACOKINETICS, SAFETY, AND EFFICACY OF 8 mg/kg OR 4 mg/kg INTRAVENOUS TOCILIZUMAB IN PATIENTS WITH MODERATE TO SEVERE COVID-19 PNEUMONIA**
PROTOCOL NUMBER:	CA42481
VERSION NUMBER:	2
IND NUMBER:	148225
NCT NUMBER:	*NCT04363736*
TEST PRODUCT:	Tocilizumab (RO4877533)
PHASE:	Phase II
INDICATION:	Moderate to severe COVID-19 pneumonia
SPONSOR:	F. Hoffmann-La Roche Ltd

Objectives and Endpoints

This Phase II study will investigate the pharmacodynamics, pharmacokinetics, safety, and efficacy of tocilizumab (TCZ) at two different doses in combination with standard-of-care (SOC) treatment in hospitalized patients with moderate to severe COVID-19 pneumonia. Specific objectives and corresponding endpoints for the study are outlined below.

Pharmacodynamic Objective

The *pharmacodynamic (PD) objective* for this study is to *assess* the *differences between two* doses of TCZ in combination with SOC for the treatment of moderate to severe COVID-19 pneumonia on the basis of the following endpoints:

- Serum concentrations of interleukin-6 (IL-6), soluble interleukin-6 receptor (sIL-6R), ferritin, and C-reactive protein (CRP) following administration of 8 mg/kg and 4 mg/kg IV TCZ at specified timepoints

Pharmacokinetic Objective

The pharmacokinetic (PK) objective for this study is to *characterize* the pharmacokinetics of two different doses of TCZ in combination with SOC for the treatment of moderate to severe COVID-19 pneumonia on the basis of the following endpoints:

- Serum concentrations of TCZ following administration of 8 mg/kg and 4 mg/kg IV TCZ at specified timepoints

Safety Objective

The safety objective for this study is to compare the safety of two different doses of TCZ in combination with SOC for the treatment of moderate to severe COVID-19 pneumonia on the basis of the following endpoints:

- Incidence and severity of adverse events, with severity determined according to National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0 (NCI CTCAE v5.0)

- *Severe acute respiratory syndrome corona virus 2* (SARS-CoV-2) (COVID-19) viral load over time, as collected by nasopharyngeal swab and bronchoalveolar lavage (BAL) samples (if applicable)
- Time to real-time polymerase chain reaction (RT-PCR) virus negativity
- The proportion of patients with any posttreatment infection
- Change from baseline in targeted clinical laboratory test results

Exploratory Objective

The exploratory objective for this study is to *assess* the efficacy of two different doses of TCZ in combination with SOC for the treatment of moderate to severe COVID-19 pneumonia on the basis of the following *endpoints*:

- Clinical status, as assessed using a 7-category ordinal scale at Days *7*, 14, and 28
- Time to clinical improvement, defined as a National Early Warning Score 2 (NEWS2) *score* of ≤ 2 maintained for 24 hours
- Time to improvement in at least two categories relative to baseline on a 7-category ordinal scale of clinical status
- Duration of supplemental oxygen use
- Incidence of intensive care unit (ICU) stays
- Incidence of mechanical ventilation
- Duration of ICU stay
- Time to clinical failure, defined as the time to death, mechanical ventilation, ICU admission, or withdrawal (whichever occurs first)

  For patients entering the study already in an ICU, clinical failure is defined as a one-category worsening on the ordinal scale, withdrawal, or death.
- Mortality rate at Days 7, 14, 21, 28, and 60
- Time to hospital discharge or "ready for discharge" (as evidenced by normal body temperature and respiratory rate, and stable oxygen saturation on ambient air or ≤ 2 L supplemental oxygen)
- Frequency of the addition of a second dose of TCZ (8 mg/kg or 4 mg/kg) at the discretion of the treating physician
- Ventilator-free days from baseline to Day 28
- Organ failure–free days from baseline to Day 28

Exploratory Biomarker Objective

The exploratory biomarker objective for this study is to identify and/or evaluate biomarkers that could provide further evidence of TCZ pharmacologic activity (i.e., mechanism of action), on the basis of the following *endpoints*:

- Assessments of *exploratory* biomarkers in relation to efficacy, exposure, and in both blood and tissue-derived samples
- *Association between TCZ pharmacokinetics and the following PD biomarkers: CRP, IL-6, sIL-6R, and ferritin*

Study Design

Description of Study

This is a Phase II, open-label, randomized, multicenter study to assess the pharmacodynamics, pharmacokinetics, safety and efficacy of two different doses of TCZ in combination with SOC in hospitalized adult patients with moderate to severe COVID-19 pneumonia. The Sponsor intends to enroll approximately 100 patients who have been diagnosed with COVID-19 pneumonia who meet the eligibility criteria at U.S. sites.

Patients must be at least 18 years old with confirmed SARS-CoV-2 (COVID-19) pneumonia per the World Health Organization criteria, including a positive PCR of any specimen (e.g., respiratory, blood, urine, stool, *or* other bodily fluids). At the time of enrollment, patients must be hospitalized and have respiratory symptoms and radiologic findings of pneumonia. Patients with severe COVID-19 pneumonia must have $SpO_2 \leq 93\%$ or $PaO_2/FiO_2 < 300$ mmHg despite receiving SOC treatment, which may include anti-viral treatment, low-dose steroids, and supportive care. Patients with moderate COVID-19 pneumonia are hospitalized patients not meeting the criteria of severe COVID-19 pneumonia and with CRP $> 2 \times$ the upper limit of normal (ULN). For both *patients with* moderate and severe COVID-19, the onset of patient symptoms must be at least 3 days prior to enrollment.

Patients who are on a mechanical ventilator $>$ 24 hours or extracorporeal membrane oxygenation (ECMO), in shock, or a combination thereof with other organ failure requiring treatment in an ICU will be excluded from the study. Patients for whom, in the opinion of the treating physician, progression to death is imminent and inevitable within the next 24 hours, irrespective of the provision of treatments, will be excluded from the study. Patients with active TB or suspected active bacterial, fungal, viral, or other infection (other than COVID-19 pneumonia) will be excluded from the study.

Patients will be randomized as soon as possible after screening in a 1:1 ratio to receive treatment with either 8 mg/kg or 4 mg/kg IV TCZ. Study treatment must be given in combination with SOC treatment.

Patients assigned to the 8-mg/kg TCZ arm will receive one IV infusion of 8 mg/kg, with a maximum dose of 800 mg, and patients assigned to the 4-mg/kg TCZ arm will receive one IV infusion of 4 mg/kg TCZ. In addition, patients in both arms will receive SOC treatment.

For patients in both arms, if a patient *has a sustained fever or clinically significant worsening of signs or symptoms (e.g., an increased supplemental oxygen requirement*), one additional infusion of TCZ at the same dose as the initial infusion can be given within 8 to 24 hours after the initial *TCZ* infusion.

Patients who do not meet the criteria for participation in this study (screen failure) may qualify for one re-screening opportunity (for a total of two screenings per participant) at the investigator's discretion. Patients are not required to re-sign the consent form if they are re-screened within 7 days after previously signing the consent form.

The study assessments to be conducted include the following: complete and limited physical examinations, measurement of vital signs and oxygen saturation, assessment of consciousness, presence and absence of respiratory support, adverse events, concomitant therapies, clinical laboratory tests, and nasopharyngeal swabs.

Patients will be followed for a total of 60 days after the first dose of study drug.

If patients are discharged from the hospital prior to Day 28, follow-up visits should occur twice weekly through Day 28. Patients are encouraged to return for onsite visits at least weekly if possible; other follow-up visits may be conducted by telephone or home visits. Patients should return to the site for the Day 28 visit. After Day 28, all patients should have follow-up visits on Days 35, 45, and 60; the Day 35 and 45 visits may be conducted by telephone or by home visits for *patients who are* discharged, while the Day 60 visit should be conducted onsite.

During the study, standard supportive care will be given according to clinical practice.

Number of Patients

Approximately 100 patients will be enrolled in this study.

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Signed Informed Consent Form by any patient capable of giving consent, or, when the patient is not capable of giving consent, by his or her legally authorized representative
- Age ≥ 18 years at time of signing Informed Consent Form
- Ability to comply with the study protocol
- Hospitalization with COVID-19 pneumonia confirmed by a positive PCR of any specimen (e.g., respiratory, blood, urine, stool, *or* other bodily fluids) and evidence of pneumonia on chest X-ray or computed tomography scan

  The onset of patient symptoms must be at least 3 days prior to enrollment.
- For severe patients, $SpO_2 \leq 93\%$ or $PaO_2/FiO_2 < 300$ mmHg

  If a patient is on supplemental oxygen with $SpO_2 > 93\%$, but desaturation to $\leq 93\%$ on lower supplemental oxygen or ambient air is documented during screening, the inclusion criterion is met.
- For moderate patients (those who do not qualify as severe based oxygen requirements), CRP $> 2 \times$ ULN is required
- For women of childbearing potential: agreement to remain abstinent (refrain from heterosexual intercourse) or use contraception, as defined below:

  Women must remain abstinent or use contraceptive methods with a failure rate of $<1\%$ per year during the treatment period and for 90 days after the final dose of TCZ. Women must refrain from donating eggs during this same period.

  A woman is considered to be of childbearing potential if she is postmenarchal, has not reached a postmenopausal state (≥ 12 continuous months of amenorrhea with no identified cause other than menopause), and is not permanently infertile due to surgery (i.e., removal of ovaries, fallopian tubes, and/or uterus) or another cause as determined by the investigator (e.g., Müllerian agenesis). The definition of childbearing potential may be adapted for alignment with local guidelines or regulations.

  Examples of contraceptive methods with a failure rate of $< 1\%$ per year include bilateral tubal ligation, male sterilization, hormonal contraceptives that inhibit ovulation, hormone-releasing intrauterine devices, and copper intrauterine devices.

  The reliability of sexual abstinence should be evaluated in relation to the duration of the clinical trial and the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not adequate methods of contraception. If required per local guidelines or regulations, locally recognized adequate methods of contraception and information about the reliability of abstinence will be described in the local Informed Consent Form.
- For men: agreement to remain abstinent (refrain from heterosexual intercourse) or use a condom, and agreement to refrain from donating sperm, as defined below:

  With a female partner of childbearing potential or pregnant female partner, men must remain abstinent or use a condom during the treatment period and for 60 days after the final dose of TCZ to avoid exposing the embryo. Men must refrain from donating sperm during this same period.

  The reliability of sexual abstinence should be evaluated in relation to the duration of the clinical trial and the preferred and usual lifestyle of the patient. Periodic abstinence (e.g., calendar, ovulation, symptothermal, or postovulation methods) and withdrawal are not adequate methods of preventing drug exposure. If required per local guidelines or regulations, information about the reliability of abstinence will be described in the local Informed Consent Form.

Exclusion Criteria

Patients who meet any of the following criteria will be excluded from study entry:

- Known severe allergic reactions to TCZ or other monoclonal antibodies
- Active TB infection
- Suspected active bacterial, fungal, viral, or other infection (besides SARS-CoV-2)
- Patients who *meet any of the following at the time of randomization*: on a mechanical ventilator $>$ 24 hours, *on* ECMO, in shock, or with other organ failure requiring treatment in an ICU
- In the opinion of the investigator, progression to death is imminent and inevitable within the next 24 hours, irrespective of the provision of treatments
- Receipt of oral anti-rejection or immunomodulatory drugs (including TCZ) within the past 3 months
- Participation in other drug clinical trials (participation in COVID-19 anti-viral trials may be permitted if confirmed by the Medical Monitor)
- ALT or AST $>10\times$ ULN detected within 24 hours at screening or at baseline (according to local laboratory reference ranges)
- ANC $<1000/\mu$L at screening and baseline (according to local laboratory reference ranges)
- Platelet count $<50{,}000/\mu$L at screening and baseline (according to local laboratory reference ranges)
- Pregnancy or breastfeeding, or positive pregnancy test at a predose examination
- Treatment with an investigational drug within 5 drug-elimination half-lives or 30 days (whichever is longer) of randomization (investigational SARS-CoV-2 [COVID-19]) anti-viral agents may be permitted after consultation with the Medical Monitor)
- Any serious medical condition or abnormality of clinical laboratory tests that, in the investigator's judgment, precludes the patient's safe participation in and completion of the study

End of Study

The end of this study is defined as the date when the last patient, last visit occurs or the date at which the last datapoint required for the last follow-up is received from the last patient, whichever occurs later. The end of the study is expected to occur approximately 2 months after the last patient is enrolled.

In addition, the Sponsor may decide to terminate the study at any time.

Length of Study

The total length of the study, from screening of the first patient to the end of the study, is expected to be approximately 4 months.

Investigational Medicinal Products

The investigational medicinal product for this study is TCZ.

Test Products (Investigational Drugs)

Patients assigned to the 4-mg/kg arm will receive the 4-mg/kg TCZ dose by IV infusion. The dose of TCZ infusion will be calculated on the basis of body weight measured prior to infusion. One additional infusion of TCZ can be given within 8 to 24 hours after the initial infusion.

Comparator

Patients assigned to the 8-mg/kg arm will receive the 8-mg/kg TCZ dose by IV infusion. The dose of TCZ infusion will be calculated on the basis of body weight measured prior to infusion. One additional infusion of TCZ can be given within 8 to 24 hours after the initial infusion.

Statistical Methods

PD Analysis

The PD outcome measures for this study are serum IL-6, sIL-6R, ferritin, and CRP levels at baseline and at specified timepoints after initiation of study drug. Data for all PD biomarkers will be presented using descriptive summary statistics, including the mean, median, range, standard deviation (SD), and coefficient of variation.

PK Analyses

The PK analysis population will consist of patients with sufficient data to enable estimation of key parameters (e.g., area under the concentration–curve [AUC], maximum serum concentration observed [C_{max}]), with patients grouped according to treatment received.

Non-linear mixed effects modeling will be used to analyze the serum TCZ concentration over time data collected in this study using existing population-PK models. Individual and mean serum TCZ concentration versus time data will be tabulated and plotted by dose level. The serum pharmacokinetics of TCZ will be summarized by estimating the total exposure (AUC), C_{max}, total clearance, and volume of distribution. Estimates for these parameters will be tabulated and summarized (mean, SD, co-efficient of variation, median, minimum, and maximum). Interpatient variability will be evaluated.

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

Determination of Sample Size

The purpose of this open-label study is the comparison of TCZ pharmacokinetics and PK-PD relationship between two doses of TCZ (8 and 4 mg/kg) in hospitalized patients with moderate to severe COVID-19 pneumonia. The comparison will be assessed using modeling and evaluating the safety and efficacy measures between Day 1 and Day 28 postrandomization. Data from the follow-up period from Days 28 to 60 may also be assessed. Approximately 100 patients will be enrolled and randomized in a 1:1 ratio to the 8-mg/kg and 4-mg/kg IV TCZ arms in this study. Disease severity type will also be stratified into moderate and severe COVID pneumonia, giving an overall equal four-way split between the two doses and type of disease severity and resulting in four groups of approximately 25 patients each. Patients may receive a repeat dose within 8 to 24 hours after the first TCZ dose if clinicians judge it is necessary (resulting in a TCZ dose range from 4 to 16 mg/kg, with a maximum dose of 800 mg).
The variability of PK parameters of TCZ in a heterogeneous COVID-19 population in an acute setting is unknown; however, 20 to 30 patients is deemed adequate considering the between-subject variability in health volunteers and patients with RA provided by population-PK modeling. Data quality issues, patient dropouts, and early hospital discharge may all lead to missing data, especially in a pandemic setting; thus, enrolling 50 patients according to dose and disease severity type (25 per arm) may offset missing data and allow an adequate volume of PK and PD data to inform the population-PK model and allow a comparison of the 4- and 8-mg/kg doses.

Interim Analysis

No interim analysis is planned for this study.

Appendix 2
Schedule of Activities: Days 1 and 2

Time after Initial Treatment (Assessment Window)	Screening [a]	Day 1		Day 2
		Baseline		
	Days −2 to 0	0 Predose (−4 hr)	15 min Postinfusion (+1 hr)	24 hr (± 4 hr)
Informed consent	x			
Review of inclusion and exclusion criteria	x	x		
Demographics	x			
Randomization		x		
Medical history		x		
Complete physical examination [c]	x			
Weight		x		
COVID-19 diagnosis [d]	x			
Chest X-ray or CT scan	x			
ECG	x			
Pregnancy test [f]	x			
PaO_2/FiO_2 [g]	x	Optional		
SpO_2 [h]	x	x	x	x
Vital signs [h]	x	x	x	x
Ordinal scoring [i]		x		x
Adverse events [j]		x		x

Appendix 2
Schedule of Activities: Days 1 and 2 (cont.)

Time after Initial Treatment (Assessment Window)	Screening [a]	Day 1		Day 2
		Baseline		
	Days −2 to 0	0 Predose (−4 hr)	15 minutes Postinfusion (+1 hr)	24 hours (± 4 hr)
Concomitant medications [k]		x		*x*
Hematology [l]	x (24 hr) [m]	x		*x*
Chemistry panel [n]	x (24 hr) [m]	x		*x*
Study drug administration [o]		x		
Central laboratory assessments				
Serum PD sample (CRP, IL-6, and sIL-6R) [p]		x [p]	x [p]	x
Serum PK sample [q]		x [q]	x [q]	x
Serum *and plasma samples* for exploratory biomarkers		x		x
SARS-CoV-2 viral load [r]		x		x
Serum SARS-CoV-2 antibody titer		x		
Cryopreserved PBMCs [s]		x		x
Blood in PAXgene® tubes for RNA analyses [t]		x		

Appendix 2
Schedule of Activities: Days 1 and 2 (cont.)

CRP = C-reactive protein; CT = computed tomography; ECG = electrocardiogram; eCRF = electronic Case Report Form; IL-6 = interleukin-6; *NEWS2 = National Early Warning Score 2*; PaO_2/FiO_2 = arterial oxygen partial pressure/fraction of inspired oxygen; PBMC = peripheral blood mononuclear cell; PCR = polymerase chain reaction; PD = pharmacodynamic; PK = pharmacokinetic; SARS-CoV-2 = severe acute respiratory syndrome corona virus (strain) 2; sIL-6R = soluble interleukin-6 receptor; SpO_2 = peripheral capillary oxygen saturation; TCZ = tocilizumab

Note: On treatment days, all assessments should be performed prior to dosing, unless otherwise specified.

[a] Results from standard-of-care tests or examinations performed prior to obtaining informed consent and within 48 hours before randomization may be used; such tests do not need to be repeated for screening. *An exception to this is clinical chemistry and hematology tests obtained prior to informed consent, which do not need to be repeated for screening if conducted within **24 hours** before randomization.*

[b] Informed consent must be documented before any study-specific screening procedure is performed.

[c] A complete physical examination, performed at screening and per the investigator's discretion during the study, includes an evaluation of the head, eyes, ears, nose, and throat, and the cardiovascular, dermatologic, musculoskeletal, respiratory, gastrointestinal, genitourinary, and neurologic systems. Any abnormality identified at screening should be recorded on the General Medical History and Baseline Conditions eCRF. New or worsened clinically significant abnormalities identified during the study should be reported as adverse events.

[d] COVID-19 test (SARS-CoV2 PCR) to confirm diagnosis should be performed within 7 days of randomization.

[e] Screening chest X-ray or CT scans should be performed within 48 hours prior to randomization. If additional chest X-rays or CT scans are taken per local practice during the study, this information should be provided on the eCRF.

[f] For women of childbearing potential, including those who have had a tubal ligation, positive urine test results will be confirmed with a serum pregnancy test. Study drug infusion must not be administered unless the serum pregnancy test result is negative.

[g] If arterial blood gases are measured.

[h] All vital sign measurements (i.e., respiratory rate, pulse rate, systolic and diastolic blood pressures, and body temperature), oxygen saturation and NEWS2-specific assessments (i.e., consciousness and presence or absence of oxygen support) should be recorded together twice daily with approximately 12 hours in between while the patient remains hospitalized. If measured more than once during a 12-hour period, the worst values (highest temperature, respiratory rate, and heart rate; lowest blood pressure, oxygen saturation, and consciousness level) during that period should be recorded on the eCRF.

[i] Assessment of clinical status using the ordinal scale should be recorded at baseline (Day 1, Visit 1) and then again daily every morning (between 8 a.m. and 12 p.m.) for patients who remain hospitalized.

[j] After informed consent has been obtained but prior to initiation of study drug, only serious adverse events caused by a protocol-mandated intervention should be reported. After initiation of study drug, all adverse events will be reported until 60 days after the final dose of study drug.

Appendix 2
Schedule of Activities: Days 1 and 2 (cont.)

[j] After this period, the Sponsor should be notified if the investigator becomes aware of any serious adverse event that is believed to be related to prior study drug treatment (see Section 5.6).

k Medication (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient in addition to protocol-mandated treatment from 7 days prior to initiation of study drug to the study completion/discontinuation visit.

l Hematology includes WBC count, RBC count, hemoglobin, hematocrit, platelet count, differential count (neutrophils, eosinophils, basophils, monocytes, lymphocytes), and lymphocyte subsets (T cells, B cells, and natural killer cells; if the test is available at the site).

m Hematology and chemistry panel assessments performed as part of standard of care, prior to consent but within 24 hours of randomization, do not need to be repeated.

n Chemistry panel (serum or plasma) includes bicarbonate or total carbon dioxide (if considered standard of care for the region), sodium, potassium, chloride, glucose, BUN or urea, creatinine, total protein, albumin, phosphorus, calcium, total bilirubin, ALP, ALT, and/or AST, uric acid, LDH, CRP, ferritin, and D-dimer.

o Study drug should be administered after collection of all samples for PK, PD, and exploratory biomarkers. The initial study drug infusion should be given within 4 hours of randomization. For patients in both arms of the study, if a patient has a sustained fever or clinically significant worsening of signs or symptoms (e.g., an increased supplemental oxygen requirement), one additional infusion of TCZ at the same dose as the initial infusion can be given within 8 to 24 hours after the initial TCZ infusion.

p On Day 1, CRP, IL-6, and sIL-6R samples should be drawn 0–4 hours before the start of infusion and then within 15 minutes (to 1 hour) after the end of the infusion from the opposite arm as the infusion. Patents receiving a second infusion of study drug should provide extra samples for CRP, IL-6, and sIL-6R prior to and 15 minutes after the end of the infusion from the opposite arm as the infusion. If the second infusion falls within the window for the Day 2 PD sample (20–24 hours) for the first infusion, it can be combined with the post infusion sample for the second infusion.

q On Day 1, PK samples should be drawn 0–4 hours before the start of infusion and then within 15 minutes (to 1 hour) after the end of the infusion from the opposite arm as the infusion. Patents receiving a second infusion of study drug should provide an extra PK sample prior to and 15 minutes (to 1 hour) after the end of the infusion from the opposite arm as the infusion. If the second infusion falls within the window for the Day 2 PK sample (20–24 hours) for the first infusion, it can be combined with the post infusion sample for the second infusion.

r Viral load will be assessed by nasopharyngeal swab. Patients who are intubated and undergo bronchoalveolar lavage will have samples taken for virological assessment. It may only be possible to access one nostril if a patient has a nasogastric tube in place, in which case sites may collect a sample from one nostril only and when possible the same nostril should be used.

s For sites capable of sample collection for analysis of T cells by high-dimensional cytometry.

t The PAXgene® blood tube should be the last tube drawn in the phlebotomy procedure to avoid contact with the RNA preservation reagent inside the tube.

Appendix 3
Schedule of Activities: Days 3–28

	Days 3–28 [a]																										*Early Tx or Study* Discon
Study Day	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
Chest X-ray or CT scan					x							x							x							x	x
Vital signs [b]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
PaO_2/FiO_2 [c]	← Optional →																										Optional
SpO_2 [b]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Ordinal scoring [d]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Adverse events [e]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Concomitant medications [f]	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Hematology [g]	x				x			x				x							x							x	x
Chemistry profile [h]	x				x			x				x							x							x	x
Central laboratory assessments																											
Serum PD sample (CRP, IL-6, and sIL-6R)	x				x							x							x							x	x
Serum PK sample	*x*				x							x							x							x	x
Serum *and plasma samples* for exploratory biomarkers	x				x							x							x							x	x
SARS-CoV-2 viral load [i]	x	x	x	x	x			x				x							x							x	x
Serum SARS-CoV-2 antibody titer																										x	x
Cryopreserved PBMCs [j]	x				x							x							x							x	x

Appendix 3
Schedule of Activities: Days 3–28 (cont.)

	Days 3–28 [a]																									*Early Tx or Study* Discon
Study Day	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
Blood in PAXgene® tubes for RNA analyses [k]	x				x																					x	x

CRP = C-reactive protein; CT = computed tomography; Discon = discontinuation; ECG = electrocardiogram; eCRF = electronic Case Report Form; IL-6 = interleukin-6; NEWS2 = National Early Warning Score; PaO_2/FiO_2 = arterial oxygen partial pressure/fraction of inspired oxygen; PBMC = peripheral blood mononuclear cell; PD = pharmacodynamic; PK = pharmacokinetic; SARS-CoV-2 = severe acute respiratory syndrome corona virus (strain) 2; sIL-6R = soluble interleukin-6 receptor; SpO_2 = peripheral capillary oxygen saturation; *Tx = treatment;* TCZ = tocilizumab.

Note: For patients who have been discharged, all assessments should be performed within ±3 days of the scheduled onsite visit.

[a] If patients are discharged from hospital prior to Day 28, follow-up visits should occur twice weekly through Day 28. Patients are encouraged to return for onsite visits at least weekly if possible; other follow-up visits may be conducted by telephone or home visits. Patients should return to the site for the Day 28 visit.

[b] All vital sign measurements (i.e., respiratory rate, pulse rate, systolic and diastolic blood pressures, and body temperature), oxygen saturation and NEWS2-specific assessments (i.e., consciousness and presence or absence of oxygen support) should be recorded together twice daily with approximately 12 hours in between while the patient remains hospitalized. If measured more than once during a 12-hour period, the worst values (highest temperature, respiratory rate, and heart rate; lowest blood pressure, oxygen saturation, and consciousness level) during that period should be recorded on the eCRF. Following hospital discharge these parameters should be recorded once at each return visit to the clinic. Vital signs and oxygen saturation will not be recorded if follow-up visits are conducted by telephone.

[c] If arterial blood gases are measured.

[d] Assessment of clinical status using the ordinal scale should be recorded at baseline (Day 1, Visit 1) and then again daily every morning (between 8 a.m. and 12 p.m.) for patients who remain hospitalized.

[e] After informed consent has been obtained but prior to initiation of study drug, only serious adverse events caused by a protocol-mandated intervention should be reported. After initiation of study drug, all adverse events will be reported until 60 days after the final dose of study drug. After this period, the Sponsor should be notified if the investigator becomes aware of any serious adverse event that is believed to be related to prior study drug treatment (see Section 5.6).

Appendix 3
Schedule of Activities: Days 3–28 (cont.)

[f] Medication (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient in addition to protocol-mandated treatment from 7 days prior to initiation of study drug to the study completion/discontinuation visit.

[g] Hematology includes WBC count, RBC count, hemoglobin, hematocrit, platelet count, differential count (neutrophils, eosinophils, basophils, monocytes, lymphocytes), and lymphocyte subsets (T cells, B cells, and natural killer cells; if the test is available at the site).

[h] Chemistry panel (serum or plasma) includes bicarbonate or total carbon dioxide (if considered standard of care for the region), sodium, potassium, chloride, glucose, BUN or urea, creatinine, total protein, albumin, phosphorus, calcium, total bilirubin, ALP, ALT, and/or AST, uric acid, LDH, ferritin, and D-dimer.

[i] Viral load will be assessed by nasopharyngeal swab. Patients who are intubated and undergo bronchoalveolar lavage will have samples taken for virological assessment. It may only be possible to access one nostril if a patient has a nasogastric tube in place, in which case sites may collect a sample from one nostril only and when possible the same nostril should be used.

[j] For sites capable of sample collection for analysis of T cells by high-dimensional cytometry.

[k] The PAXgene® blood tube should be the last tube drawn in the phlebotomy procedure to avoid contact with the RNA preservation reagent inside the tube

Appendix 4 Schedule of Activities: After Day 28

Study Day (Assessment Window)			Study Completion
	35 [a] (±3 days)	45 [a] (±3 days)	60 (±3 days)
Chest X-ray or CT scan			x
SARS-CoV-2 viral load [b]	x	x	x
Vital signs [c]	x	x	x
SpO_2 [c]	x	x	x
Ordinal scoring [d]	x	x	x
Adverse events [e]	x	x	x
Concomitant medications [f]	x	x	x
Hematology [g]	x	x	x
Chemistry panel [h]	x	x	x
Central laboratory assessments			
Serum PD sample (CRP, IL-6, and sIL-6R)	x		x
Serum PK sample	x		x
Serum *and plasma samples* for exploratory biomarkers	x		x
SARS-CoV-2 viral load [i]			*x*
Cryopreserved PBMCs [j]			*x*
Blood in PAXgene® tubes for RNA analyses [k]			*x*
Serum SARS-Cov-2 antibody titer			x

CRP = C-reactive protein; CT = computed tomography; eCRF = electronic Case Report Form; IL-6 = interleukin-6; PD = pharmacodynamic; PK = pharmacokinetic; SARS-CoV-2 = severe acute respiratory syndrome corona virus (strain) 2; sIL-6R = soluble interleukin-6 receptor; SpO_2 = peripheral capillary oxygen saturation.

Appendix 4 Schedule of Activities: After Day 28 (cont.)

[a] If patients are unable to return for onsite visits at Day 35 and/or Day 45, these may be conducted by telephone or home visits. Patients should return to the site for the Day 60 study completion visit.

[b] Patients who remain in hospital will have viral load assessed by nasopharyngeal swabs; these will be done if there is evidence of ongoing infection.

[c] For patients who remain in the hospital, vital sign measurements and NEWS2-specific assessments should be conducted twice daily. Following hospital discharge, these parameters should be recorded once at each return visit to the clinic. Vital signs and oxygen saturation will not be recorded if follow-up visits are conducted by telephone.

[d] Assessment of clinical status using the ordinal scale should be recorded at baseline (Day 1, Visit 1)δ and then again daily every morning (between 8 a.m. and 12 p.m.) for patients who remain hospitalized.

[e] After informed consent has been obtained but prior to initiation of study drug, only serious adverse events caused by a protocol-mandated intervention should be reported. After initiation of study drug, all adverse events will be reported until 60 days after the final dose of study drug. After this period, the Sponsor should be notified if the investigator becomes aware of any serious adverse event that is believed to be related to prior study drug treatment (see Section 5.6).

[f] Medication (e.g., prescription drugs, over-the-counter drugs, vaccines, herbal or homeopathic remedies, nutritional supplements) used by a patient in addition to protocol-mandated treatment from 7 days prior to initiation of study drug to the study completion/discontinuation visit.

[g] Hematology includes WBC count, RBC count, hemoglobin, hematocrit, platelet count, differential count (neutrophils, eosinophils, basophils, monocytes, lymphocytes), and lymphocyte subsets (T cells, B cells, and natural killer cells). Hematology labs will not be performed if follow-up visits are conducted by telephone.

[h] Chemistry panel (serum or plasma) includes bicarbonate or total carbon dioxide (if considered standard of care for the region), sodium, potassium, chloride, glucose, BUN or urea, creatinine, total protein, albumin, phosphorus, calcium, total bilirubin, ALP, ALT, and/or AST, uric acid, LDH, ferritin, and D-dimer. Chemistry laboratory assessments will not be performed if follow-up visits are conducted by telephone.

[i] *Viral load will be assessed by nasopharyngeal swab. Patients who are intubated and undergo bronchoalveolar lavage will have samples taken for virological assessment. It may only be possible to access one nostril if a patient has a nasogastric tube in place, in which case sites may collect a sample from one nostril only and when possible the same nostril should be used.*

[j] *For sites capable of sample collection for analysis of T cells by high-dimensional cytometry.*

[k] *The PAXgene® blood tube should be the last tube drawn in the phlebotomy procedure to avoid contact with the RNA preservation reagent inside the tube.*

Appendix 5
National Early Warning Score 2 (NEWS2)

Physiological parameter	Score 3	2	1	0	1	2	3
Respiration rate (per minute)	≤8		9–11	12–20		21–24	≥25
SpO_2 Scale 1 (%)	≤91	92–93	94–95	≥96			
SpO_2 Scale 2 (%)	≤83	84–85	86–87	88–92 ≥93 on air	93–94 on oxygen	95–96 on oxygen	≥97 on oxygen
Air or oxygen?		Oxygen		Air			
Systolic blood pressure (mmHg)	≤90	91–100	101–110	111–219			≥220
Pulse (per minute)	≤40		41–50	51–90	91–110	111–130	≥131
Consciousness				Alert			CVPU
Temperature (°C)	≤35.0		35.1–36.0	36.1–38.0	38.1–39.0	≥39.1	

SpO_2 = oxygen saturation.

Oxygen saturation (SpO_2) should be scored according to either the SpO_2 Scale 1 or 2 presented in the table above. The SpO_2 Scale 2 is for patients with a target oxygen saturation requirement of 88%–92% (e.g., in patients with hypercapnic respiratory failure related to advanced lung diseases, such as chronic obstructive pulmonary disease). This should only be used in patients confirmed to have hypercapnic respiratory failure by blood gas analysis on either a prior or their current hospital admission.

The decision to use the SpO_2 Scale 2 should be made by the treating physician and should be recorded in the eCRF. In all other circumstances, the SpO_2 Scale 1 should be used.

For physiological parameter “Air or Oxygen?”: Any patients requiring the use of oxygen or other forms of ventilation to maintain oxygen saturations and support respiration should be assigned a score of 2.

The consciousness level should be recorded according to the best clinical condition of the patient during the assessment. Patients who are assessed as “Alert” (A) should be assigned a score of 0. Patients assessed as “New Confusion” (C), “Responsive to Voice” (V), “Responsive to Pain” (P), or “Unconscious” should be assigned a score of 3.

Scores should be assigned for respiratory rate, systolic blood pressure, pulse, and temperature according to the table above.

Appendix 5
National Early Warning Score 2 (NEWS2) (cont.)

NEWS2 values will be calculated electronically throughout the study by the Sponsor based upon entry of vital sign parameters by the investigator on the appropriate electronic Case Report Form.

Example Case Calculation:

An 82-year-old woman was admitted, tested positive to COVID-19, and admitted to high-dependency unit for non-invasive ventilation. Her taken observations and corresponding NEWS2 score are as follows:

Physiological Parameter	Observation	Component Score
Respiratory rate (per minute)	26	3
Oxygen saturation (SpO_2 %)	95%	1
Supplemental oxygen	Yes	2
Systolic blood pressure (mmHg)	95	2
Pulse rate (bpm)	109	1
Conscious level	New confusion	3
Temperature (^oC)	39	1
	Total NEWS2 score	**13**

REFERENCE

Royal College of Physicians. National early warning score (NEWS) 2. Standardizing the assessment of acute-illness severity in the NHS. London: RCP, 2017.

Appendix 6
Cochran-Mantel-Haenszel Test

- The weighted difference in proportions is the difference in the response rates in the experimental treatment group compared with the control treatment group, adjusted for any stratification factors. With two stratification factors, the number of patients in each strata is defined as nijk where i is the level of the first stratification factor and j is the level of the second stratification factor and k is treatment group (experimental or control). The number of events in each strata is denoted by xijk, where i, j and k are as above. The proportion of responders in each strata will be calculated by:

$$p_{ijk} = \frac{x_{ijk}}{n_{ijk}}$$ where i, j and k are as above

- The difference in proportions for each strata will then be calculated as the proportion of patients in each strata in the experimental treatment group (EXP) minus the proportion of patients in each strata in the control treatment group (CON) and denoted $d_{ij} = p_{ijEXP} - p_{ijCON}$, for i and j as above.
- The weights for each strata (i, j) will be calculated as follows:

$$w_{ij} = \frac{n_{ijEXP} * n_{ijCON}}{n_{ijEXP} + n_{ijCON}}$$

- Within each strata, the weighted differences in the proportions in each of the treatment groups will be calculated as follows:

$$wd_{ij} = w_{ij} d_{ij}$$

- and then summed:

$$WD = \sum_i \sum_j wd_{ij}$$

- After calculation of the weighted difference in proportions, the calculation of the 95% CI is as follows:
- Continuity-corrected Proportions

$$p^{\#}_{ijk} = \frac{x_{ijk} + 0.5}{n_{ijk} + 1}$$

- Variances

$$Upvar_{ij} = w_{ij}^2 \left[p^{\#}_{ijEXP} \frac{\left(1 - p^{\#}_{ijEXP}\right)}{n_{ijEXP}} + p^{\#}_{ijCON} \frac{\left(1 - p^{\#}_{ijCON}\right)}{n_{ijCON}} \right]$$

Appendix 6
Cochran-Mantel-Haenszel Test (cont.)

To calculate the sum of the weights and variances over all strata:

Sum over Strata

- $$W = \sum_i \sum_j w_{ij}$$ (sum of weights)

$$Var = \sum_i \sum_j Up\,\mathrm{var}_{ij}$$ (sum of variances)

Point Estimate and Standard Error

$$d = \frac{WD}{W}; \quad se = \sqrt{\frac{Var}{W^2}}$$

Stratified 95% Confidence Intervals

$$\text{Lower Limit} = d - 1.96se$$

$$\text{Upper Limit} = d + 1.96se$$

Appendix 7 Handling of Values Below the Limit of Quantification (BLQ)

PD and PK values can be below the limit of quantification (BLQ). The rules to handle those values are specified next.

If less than 50% of the PK values at a given time point are BLQ then we will follow the ½ Cmax rule:

1. Set all BLQ/LTR observations to zero if BLQ/LTR occurs prior to cmax.
2. Set all BLQ/LTR observations to missing if BLQ/LTR occurs after cmax.
3. **If > 1/2 of the values are LTR/BLQ (i.e., percentage sample with LTRs/BLQs > 50%) then, FLAGSUM = 3, otherwise missing.**

If more than 1/3 of the PD values at a given time point are BLQ then we will follow the 1/3 rule:

1. For pre-dose PK samples (i.e. nominal time <= 0) if sample result is LTR/BLQ (i.e. if PC.PCSTRESC in ('LTR', 'BLQ')) then the result is replaced as 0 (i.e. SUMMCONC = 0 and GRAFCONC=0)
2. For post-dose PD samples (i.e. nominal time > 0) then those LTR values are replaced by 1/2*LLOQ (i.e. if PC.PCSTRESC in ('LTR', 'BLQ') then SUMMCONC=0.5*LLOQ and GRAFCONC=0.5*LLOQ)
3. **If > 1/3 sample results are LTR/BLQ (i.e., percentage of sample with LTRs/BLQs > 33%) then FLAGSUM=2, otherwise missing.**

*LTR: Less Than Reportable

Appendix 8 Additional descriptions of time windows

Headings mean:
AVISIT: character representation of visit
AVISITN: numeric representation of visit
AWTARGET: target day (approximately the window midpoint)
AWLO: lower window range
AWHI: higher window range

All vital signs, including oxygen saturation and NEWS2

AVISIT	AVISITN	AWTARGET	AWLO	AWHI
PRE-BASELINE	0	1	.	1
BASELINE	1	1	.	1
Day 1 (Postdose)	1.5	1	1	1
Day 2	2	2	2	2
Day 3	3	3	3	3
Week 1	7	7	4	10
Week 2	14	14	11	17
Week 3	21	21	18	24
Week 4	28	28	25	28
Week 5	35	35	29	38
Day 45	45	45	39	52
Day 60	60	60	53	>60

Safety labs

AVISIT	AVISITN	AWTARGET	AWLO	AWHI
PRE-BASELINE	0	1	.	1
BASELINE	1	1	.	1
Day 1 (Postdose)	1.5	1	1	1
Day 2	2	2	2	2
Day 3	3	3	3	3
Day 7	7	7	4	8

Appendix 8 Additional descriptions of time windows (cont.)

Day 10	10	10	9	12
Week 2	14	14	13	17
Week 3	21	21	18	24
Week 4	28	28	25	28
Week 5	35	35	29	38
Day 45	45	45	39	52
Day 60	60	60	53	>60

PD and PK parameters

AVISIT	AVISITN	AWTARGET	AWLO	AWHI
BASELINE	1	1	.	1
Day 1 (Postdose)	1.5	1	1	1
Day 2	2	2	2	2
Day 3	3	3	3	3
Week 1	7	7	4	10
Week 2	14	14	11	17
Week 3	21	21	18	24
Week 4	28	28	25	28
Week 5	35	35	29	38
Day 60	60	60	39	>60

Death and discharge

AVISIT	AVISITN	AWTARGET	AWLO	AWHI
PRE-BASELINE	0	1	.	1
BASELINE	1	1	.	1
Week 1	7	7	1	7
Week 2	14	14	8	14
Week 3	21	21	15	21

Appendix 8 Additional descriptions of time windows (cont.)

Week 4	28	28	22	28
Week 5	35	35	29	35
Day 45	45	45	36	45
Day 60	60	60	46	>60

Ordinal scoring and all other incidence measures

AVISIT	AVISITN	AWTARGET	AWLO	AWHI
PRE-BASELINE	0	1	.	1
BASELINE	1	1	.	1
Day 2	2	2	2	2
Day 3	3	3	3	3
Week 1	7	7	4	10
Week 2	14	14	11	17
Week 3	21	21	18	24
Week 4	28	28	25	28
Week 5	35	35	29	38
Day 45	45	45	39	52
Day 60	60	60	53	>60