Official Title: A Phase II, Randomized, Double-Blind, Placebo-Controlled, Multicenter Study to Evaluate the Safety and Efficacy of MSTT1041A or UTTR1147A in Patients With Severe Covid-19 Pneumonia

NCT Number: NCT04386616

Document & Date: Statistical Analysis Plan, Version 1: 13-January-2021

STATISTICAL ANALYSIS PLAN

TITLE:	**A PHASE II, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE SAFETY AND EFFICACY OF MSTT1041A OR UTTR1147A IN PATIENTS WITH SEVERE COVID-19 PNEUMONIA**
PROTOCOL NUMBER:	GA42469
STUDY DRUG:	MSTT1041A (RO7187807) UTTR1147A (RO7021610)
VERSION NUMBER:	1
IND NUMBER:	149507
NCT NUMBER:	NCT04386616
SPONSOR:	Genentech, Inc.
PLAN PREPARED BY:	[redacted], Ph.D.
DATE FINAL:	See electronic date stamp below

STATISTICAL ANALYSIS PLAN APPROVAL

TABLE OF CONTENTS

1. BACKGROUND

This document provides statistical methods to be used in the analyses of GA42469, including primary analysis, and potential interim analysis. The analysis of biomarkers and exploratory PK/ADA objectives will not be covered in this document.

The study GA42469 is a Phase II, randomized, double-blind, placebo-controlled, multicenter study to assess the efficacy and safety of MSTT1041A or UTTR1147A in combination with SOC compared with matching placebo in combination with SOC in patients hospitalized with severe COVID-19 pneumonia.

The primary endpoint of the study is time to recovery defined as time to score of 1 or 2 on the 7-category ordinal scale (see details in Section 2.2.1).

2. STUDY DESIGN

2.1 PROTOCOL SYNOPSIS

The Protocol Synopsis is in Appendix 1. For additional details, see the Study Flowchart in Appendix 2.

2.2 OUTCOME MEASURES

2.2.1 Primary Efficacy Endpoint

The primary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo in combination with SOC on the basis of the following endpoint:

- Time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first)

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 $\leq$2 L supplemental oxygen)
2. Non-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 ECMO or mechanical ventilation and additional organ support (e.g., vasopressors, renal replacement therapy)
7. Death

2.2.2 Secondary Efficacy Endpoints

The secondary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo in combination with SOC on the basis of the following endpoints:

- Time to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical status
- 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 $\leq$ 2 L supplemental oxygen)
- Duration of supplemental oxygen
- Proportion of patients alive and free of respiratory failure (requiring non-invasive ventilation, high-flow oxygen, mechanical ventilation, or ECMO) at Day 28
- Clinical status assessed using a 7-category ordinal scale at Days 14 and 28
- Incidence of invasive mechanical ventilation or ECMO
- Ventilator-free days to Day 28
- Incidence of ICU stay
- Duration of ICU stay
- Time to clinical failure, defined as the time to death, mechanical ventilation, ICU admission, or withdrawal of care (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 rate at Days 14 and 28
- Time to clinical improvement, defined as a National Early Warning Score 2 (NEWS2) of $\leq$ 2 maintained for 24 hours

2.2.3 Exploratory Efficacy Endpoints

The exploratory efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo on the basis of the following endpoints:

- Incidence of vasopressor use
- Duration of vasopressor use
- Incidence of starting hemodialysis
- SARS-CoV-2 viral load on Day 15 and Day 28
- Proportion of patients with secondary bacterial infections

2.2.4 Pharmacokinetic Efficacy Endpoints

The PK objective for this study is to characterize the MSTT1041A and UTTR1147A PK profiles on the basis of the following endpoints:

- Serum concentration of MSTT1041A at specified timepoints

- Serum concentration of UTTR1147A at specified timepoints

2.2.5 Safety Endpoints

The safety objective for this study is to evaluate the safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo in combination with SOC for the treatment of 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)
- Change from baseline in targeted vital signs, targeted clinical laboratory test results, and targeted ECG parameters

2.3 DETERMINATION OF SAMPLE SIZE

A total of approximately 390 patients will be randomly allocated in a 2:2:1:1 ratio to receive MSTT1041A, UTTR1147A, or their matching placebos. The sample size provides approximately 80% power using a log-rank Chi-square test to detect a 7-day difference between treatment groups in time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first), under the following assumptions: median time to recovery in the placebo group is 21 days, with 28 days follow-up, and using a one-sided 5% alpha. The minimal detectable difference is expected to be approximately 5.3 days.

2.4 ANALYSIS TIMING

The primary study analysis will be conducted when the last patient has either withdrawn or completed their Day 28 visit, and will be based on cleaned data for all patients up to and inclusive of their Day 28 assessment.

At the time of the primary analysis, the Sponsor personnel who are analyzing data from the first 28 days of the study will be unblinded to treatment assignment. This includes personnel directly involved in the statistical analyses and programming activities as well as Sponsor personnel from other functions (e.g., Clinical Science, Safety Science, Clinical Pharmacology, and Regulatory Affairs) who will be involved in assessing, summarizing, and interpreting the data.

The final analysis will be conducted when all patients have completed the Day 60 study completion visit or discontinued from study early, and all data from the study are in the database, and the database is locked. A Clinical Study Report (CSR) will be produced based on the final analysis.

3. STUDY CONDUCT

3.1 RANDOMIZATION

Patients will be randomly assigned to one of four treatment arms: MSTT1041A, UTTR1147A, or their matching placebos. Randomization will occur in a 2:2:1:1 ratio

through use of a permuted-block randomization method to ensure a balanced assignment to each treatment arm. Randomization will be stratified according to need for mechanical ventilation (yes vs. no) and region (North America, South America, Western Europe, and Other). The proportion of patients with a need for mechanical ventilation will be capped at approximately 25% of the overall study population.

3.2 DATA MONITORING

A Data Monitoring Committee (DMC) will review unblinded safety and efficacy data to assess whether treatment with MSTT1041A or UTTR1147A is associated with toxicity or worsening disease. Members of the DMC will include representatives from Clinical Science and Drug Safety who are not directly involved in the study and representatives from Biostatistics and Statistical Programming and Analysis and may invite representatives from other functional areas on an ad-hoc basis when additional expertise is required (e.g., Clinical Pharmacology, Research). The DMC will also include at least two external experts in the field, and additional experts may be added during the course of the study.

The DMC will review cumulative safety data through at least Day 14 after the first 30 patients have been enrolled or after the first 30 days from the first patient enrolled (whichever occurs first). Subsequently, the DMC will review cumulative safety and efficacy data after each additional 75 patients have been enrolled or 3 months (whichever occurs first); for example, once approximately 75, 150, 225, and 300 patients have enrolled. Ad hoc meetings may be held at the request of the DMC or Sponsor at any time to address potential safety concerns. The DMC will have access to all available data at each review to perform an overall benefit/risk assessment. A detailed description of the procedures, data flow, possible recommendations and meeting schedule will be provided in the DMC Charter.

A Safety Monitoring Committee (SMC) consisting of internal team members will review blinded safety and efficacy data to assess whether treatment with MSTT1041A or UTTR1147A is associated with toxicity or worsening disease. The SMC will meet to review safety and efficacy data at least as frequently as the DMC. The SMC may also meet to assess the significance of other adverse events or safety findings at any time following a request from either an investigator or the Sponsor.

4. STATISTICAL METHODS

4.1 ANALYSIS POPULATIONS

Disposition summaries will be based on the all patients randomized into the study. Efficacy analyses will be conducted on the modified intent-to-treat (mITT) population. Safety analyses will be conducted for the safety population.

4.1.1 Modified Intent-to-Treat Population

The mITT population is defined as all patients randomized in the study who received at least one dose of study drug, with patients grouped according to the treatment assigned at randomization.

4.1.2 Safety Population

Safety population is defined as patients who received at least one dose of study drug, with patients grouped according to the treatment received.

4.2 ANALYSIS OF STUDY CONDUCT

The number of patients who enroll, discontinue, or complete the study will be summarized by treatment group. Reasons for premature study discontinuation will be listed and summarized. Enrollment and major protocol deviations will be listed and summarized.

Deviation on eligibility criteria and other major protocol deviation will be listed and summarized by treatment group.

4.3 ANALYSIS OF TREATMENT GROUP COMPARABILITY

Demographic and baseline characteristics (including, but not limited to, age, sex, BMI, days from diagnosis to randomization, geographic region, ordinal scale for clinical status, mechanical ventilation) will be summarized using means, standard deviations, medians, and ranges for continuous variables and proportions for categorical variables, as appropriate. Summaries will be presented overall and by treatment groups for the mITT and safety population.

4.4 EFFICACY ANALYSIS

Efficacy analyses will be conducted on the mITT population, consisting of all patients randomized in the study who received at least one dose of study drug, with patients grouped according to the treatment assigned at randomization.

Comparisons of Interest

MSTT1041A placebo and UTTR1147A placebo will be pooled in efficacy analyses. Comparison of efficacy will be performed between each of the two treatment groups and the pooled placebo. That is, there are two comparisons:

- MSTT1041A with pooled placebo
- UTTR1147A with pooled placebo

Summaries of demographic and baseline characteristics by MSTT1041A placebo UTTR1147A placebo will be produced to assess the comparability of the two placebo groups before pooling.

Efficacy analyses will not involve formal comparison between MSTT1041A and UTTR1147A.

Type I Error Management

Type I error will be controlled for the primary endpoint using the Hochberg procedure at the 5% significance level.

No control of type I error will be applied to secondary and exploratory endpoints, including sensitivity and biomarker analyses.

Covariate Adjustment

Unless otherwise noted, analyses of efficacy outcome measures will be adjusted by the need for mechanical ventilation [yes vs. no] and region [country].

Baseline Definition

Unless otherwise specified, baseline is defined as the last available pre-treatment value taken on or before the day of randomization for all assessments.

Baseline measurements will be used for the summary of demographic characteristics, as well as for all change-from-baseline analysis of efficacy, safety and PD endpoints.

4.4.1 Primary Efficacy Endpoint

The primary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo in combination with SOC on the basis of the following endpoint:

- Time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first)

Time to recovery will be analyzed using the stratified log-rank test, adjusting for covariates specified in Section 4.4. The Kaplan-Meier plot, median time to event, and their 95% CI, and p-value from the stratified log-rank test will be presented. A Cox proportional hazards regression model will be used to estimate the hazard ratio comparing MSTT1041A or UTTR1147A with placebo, respectively, adjusting for stratification factors. Hazard ratio and 95% CI will be presented. In addition, p-value from unstratified log-rank test, unadjusted hazard ratio, and 95% CI will also be presented.

For patients who are lost to follow-up or discontinue from study early without event, they will be censored at the day of their last observed assessment. For patients who completed study follow-up without events, they will be censored at their Day 28 visit. For patients who died without events before the Day 28, they will be right censored at Day 28. For patients with baseline ordinal score of 1 or 2, they will be censored at their Day 28 visit.

4.4.2 Secondary Efficacy Endpoints

Clinical Status Endpoints

Among secondary endpoints, clinical status endpoints include:

- Clinical status assessed using a 7-category ordinal scale at Days 14 and 28

At each time point of interest, the clinical status endpoint will be analyzed using a proportional odds model, adjusting for covariates specified in Section 4.4. The odds ratio, p-value, and 95% CI will be presented. Stacked bar charts of the ordinal scale will be produced by treatment groups.

For patients who withdraw prior to the analysis timepoint, their last known post baseline ordinal score prior to withdrawal will be used in the primary analysis, unless death within the 28-day time frame was capture from public records or otherwise, in which case death will be used in the analysis.

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. Should the proportional odds assumption be violated, a Van Elteren Test will be used to compare treatment groups.

In addition, the proportion of patients with a response in each category of the ordinal scale will be summarized by treatment groups at earlier time points including but not limited to Day 7 and Day 21.

Time to Event Endpoints

Among secondary endpoints, time to event endpoints include:

- Time to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical status
- 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 $\leq$2 L supplemental oxygen)
- Time to clinical failure, defined as the time to death, mechanical ventilation, ICU admission, or withdrawal of care (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.
- Time to clinical improvement (TTCI), defined as a National Early Warning Score 2 (NEWS2) of $\leq$2 maintained for 24 hours

Time to event endpoints will be analyzed similarly as the primary endpoint, as described in Section 4.4.1. For patients who are lost to follow-up or discontinue from study early without event, they will be censored at the day of their last observed assessment. For

each of the time to event endpoints, patients completed study follow-up without events will be censored at Day 28 visit. For time to event endpoints, other than time to clinical failure, deaths will be right censored at Day 28. For TTCI, patients discharged from hospital without clinical improvement will be censored at the day of their last observed assessment. For patients that are discharged without an ordinal score assessment at discharge, they will be assumed to have ordinal score of 1 on the day of discharge.

Incidence Endpoints

Among secondary endpoints, incidence endpoints include:

- Proportion of patients alive and free of respiratory failure (requiring non-invasive ventilation, high-flow oxygen, mechanical ventilation, or ECMO) at Day 28
- Incidence of invasive mechanical ventilation or ECMO
- Incidence of ICU stay
- Mortality rate at Days 14 and 28

Incidence endpoints will be analyzed using the Cochran-Mantel-Haenszel (CMH) test statistics, adjusting for stratification factors. For treatment group comparison between MSTT1041A or UTTR1147A and placebo, respectively, difference in proportions, 95% CI and p-value will be presented.

The incidence of mechanical ventilation and the incidence of respiratory failure will be derived from the Vital Signs and Oxygen Saturation form. The incidence of ICU stay will be derived from the ICU Stay Information Log using the number of distinct ICU admission date/time. Mortality will be derived from the availability of death dates.

For the incidence of mechanical ventilation and the incidence of ICU stay, responses within the 28-day treatment period will be considered. For all incidence endpoints, except mortality, patients who are lost to follow-up, discontinue from study, or die prior to Day 28 will be considered as non-responders. For mortality, all deaths post discontinuation or discharge will be included in the analysis, and cumulative deaths will be analyzed at the time points of interest, including but not limited to Day 14 and 28.

Duration Endpoints

Among secondary endpoints, duration endpoints include:

- Duration of supplemental oxygen
- Ventilator-free days to Day 28
- Duration of ICU stay

Duration endpoints will be analyzed using the stratified Wilcoxon Rank Sum test, adjusting for stratification factors. The median duration for each treatment group, 95% CI

for the median, and p-value comparing MSTT1041A or UTTR1147A with placebo, respectively, will be presented.

The number of Ventilator-free days (VFDs) is defined as the number of days during the 28-day treatment period when the patient is alive and without need for invasive mechanical ventilation. VFDs will be derived from the Vital Signs and Oxygen Saturation form. For any day during Day 1 and Day 28, if invasive mechanical ventilation or ECMO is recorded for any part of the day (>= 12 hours during mechanical invasive ventilation for patients with tracheostomy), the day will not be counted as a VFD; otherwise, the day will be counted. For any days prior to Day 28 where status of mechanical ventilator is missing, the last known status will be carried forward. The total number of days will be the sum of all VFDs, regardless of whether the days occur consecutively or in non-consecutive intervals. Special considerations for calculating VFD include the following:

- For patients who are on an invasive mechanical ventilator from Day 1 to Day 28, their VFDs will be zero if they complete the study on or prior to Day 28.
- For patients who discontinue from the study early while being on invasive mechanical ventilator, their remainder of the days, i.e., from the day of discontinuation to Day 28, will not be counted as VFDs.
- For patients who discontinue from the study early without being on invasive mechanical ventilator, their remainder of the days, i.e., from the day of discontinuation to Day 28, will be counted as VFDs.
- For patients who died on or prior to Day 28, their VFDs will be zero.

Duration of ICU stay will be calculated as the total number of hours spent in ICU up to and inclusive of 28 days. ICU duration will be derived from the ICU Stay Information Log using the difference between ICU discharge date/time and ICU admission date/time. If ICU admission occurs before randomization, the ICU duration will be counted from the date of dosing. Partial admission and discharge date/time will be imputed following a conservative approach. For each patient, durations of multiple ICU stays will be summed. Special considerations for calculating ICU duration include the following:

- For patients who discontinue from study early and are in ICU on the day of discontinuation, they will be assumed to be in ICU for the remainder of the days, i.e., from the day of discontinuation to Day 28.
- For patients who discontinue from the study early and are not in ICU on the day of discontinuation, they will be assumed to have no incidence of ICU after discontinuation.
- For patients who are discharged from the hospital, any ongoing ICU stays without an ICU discharge date/time will be imputed from the date/time of hospital discharge. The discharged patients will be assumed to have no incidence of ICU stay after discharge.
- For patients who die on or prior to Day 28, their duration of ICU stay will be 28 days.

Duration of supplemental oxygen is defined as the number of days during the 28-day treatment period when the patient is alive and receives "Supplemental Oxygen or other forms of ventilation", as recorded in the Vital Signs and Oxygen Saturation form. For each patient, the duration of multiple non-consecutive periods during which the patient receives supplemental oxygen will be summed. For any days prior to Day 28 where status of supplemental oxygen use is missing, the last known status will be carried forward. Special considerations for calculating the duration of supplemental oxygen include the following:

- For patients who discontinue from study early and are on supplemental oxygen on the day of discontinuation, they will be assumed to receive supplemental oxygen for the remainder of the days, i.e., from the day of discontinuation to Day 28.
- For patients who discontinue from study early and are not on supplemental oxygen on the day of discontinuation, they will be assumed not to receive supplemental oxygen for the remainder of the days, i.e., from the day of discontinuation to Day 28.
- For patients who die on or prior to Day 28, their duration of supplemental oxygen will be 28 days.

4.4.3 Exploratory Efficacy Endpoints

Exploratory efficacy endpoints will be analyzed using the same methods as described for the secondary endpoints in Section 4.4.2 where appropriate.

- Incidence of vasopressor use
- Duration of vasopressor use
- Incidence of starting hemodialysis
- SARS-CoV-2 viral load on Day 15 and Day 28
- Proportion of patients with secondary bacterial infections

Incidence of vasopressor use will be derived from the Vital Signs and Oxygen Saturation form. SARS-CoV-2 viral load will be derived from the COVID 19 Viral Local form.

Duration of vasopressor use will be defined as the number of days during the 28-day treatment period when the patient is alive and receive vasopressor, and will be calculated similarly as duration of supplemental oxygen described in Section 4.4.2.

4.4.4 Sensitivity Analyses

For primary efficacy endpoint, sensitivity analyses to evaluate the result robustness will be conducted by comparing each of the two treatment groups with their matching placebos, respectively. That is, sensitivity analyses for the primary endpoint will be conducted with the following comparisons.

- MSTT1041A with MSTT1041A placebo

- UTTR1147A with UTTR1147A placebo

For primary endpoint and other time to event endpoints, other than time to clinical failure, sensitivity analyses may be performed using a competing risk model, either cause-specific hazard model or the sub-distribution hazard model (Fine and Gray, 1999). In competing risk analysis, deaths within the 28-day follow-up period will be coded as a competing risk at the time of death. Cumulative incidence functions will be presented for both the primary event and the competing event of death.

4.4.5 Subgroup Analyses

Exploratory subgroup analyses on mortality will be performed to evaluate result consistency across pre-specified subgroups defined by demographic and baseline characteristics. The following subgroups will be analyzed. Other subgroup analyses may also be performed.

- Age (< 65 years, >= 65 years)
- Sex (male, female)
- Race
- Region (US, ex-US)
- Mechanical ventilation at randomization (Yes, No)
- Ordinal clinical status (category 3-4, category 5-6)
- COVID-19 treatments (Steroid, Redesivir, Tocilizumab)

4.5 PHARMACOKINETIC AND PHARMACODYNAMIC ANALYSES

The PK analysis population will consist of patients who received at least one dose of MSTT1041A or UTTR1147A and have sufficient data to enable estimation of key parameters (e.g., C_{max}), with patients grouped according to treatment received.

Estimates for the PK parameters will be tabulated and summarized (mean, standard deviation, coefficient of variation, median, minimum, and maximum), as appropriate. Individual and mean serum MSTT1041A or UTTR1147A concentration versus time data will be tabulated by dose level. Additional PK analyses will be conducted as appropriate.

4.6 SAFETY ANALYSES

Safety analyses will be conducted on all patients who received at least one dose of study drug, with patients grouped according to the treatment received.

Safety will be assessed through summaries of exposure to study treatment, adverse events, laboratory test results, vital signs, and ECGs.

Relevant laboratory, vital sign, and ECGs will be summarized by appropriate descriptive statistics by treatment groups. For selected parameters, proportion of patients

experiencing clinically significant changes relative to baseline will be summarized by treatment arm.

4.6.1 Exposure of Study Medication

Exposure to study drug will be summarized as the number of patients with one or two doses and the number of patients with infusion modification by treatment groups.

4.6.2 Adverse Events

All verbatim adverse event terms will be mapped to MedDRA thesaurus terms, adverse event treatment will be mapped to WHO Drug Global B3 Format dictionary, adverse event severity will be graded according to NCI CTCAE v5.0 scale. A glossary of these codes will be produced.

A treatment-emergent adverse event is defined as any new adverse event reported or worsening of an existing condition on or after the first dose of study drug during. Only treatment-emergent adverse events will be summarized, including the following:

- All adverse events
- Serious adverse events
- Adverse events leading to death
- Adverse event of special interest
- Adverse events leading to discontinuation from study treatment
- Adverse events leading to discontinuation from study
- Deaths and causes of death

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 arms. 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.

Adverse event of special interest will be derived from the Adverse Event form, including the following:

- Case of an elevated ALT or AST 40 in combination with either an elevated bilirubin or clinical jaundice, as defined in protocol
- Suspected transmission of an infectious agent by the study drug Suspected infusion related reactions or hypersensitivity reactions, including anaphylaxis, within 24 hours after infusion
- MACE
- Grade >=3 rash

Listings of AEs and SAEs will be produced. Adverse events of special interest will also be listed.

4.6.3 Laboratory Data

Laboratory value will be converted to Système International units, and data will be transformed to a common Roche Standard Reference Range. Summary tables will be created on the actual values and changes from baseline by parameters over visits by treatment arm. The numbers and proportions of patients outside the normal upper and lower limits will also be summarized by parameters over visits by treatment arm.

4.6.4 Vital Signs

Summary tables will be created on the actual values and changes from baseline by parameters over visits by treatment arm. The numbers and proportions of patients outside the normal upper and lower limits will also be summarized by parameters over visits by treatment arm. The reference range of vital signs are as follows.

Vital Sign	Unit	Analysis Normal Range Lower Limit	Analysis Normal Range Upper Limit
Pulse Rate	beats/min	60	100
Temperature	C	36.5	38
Systolic Blood Pressure	mmHg	90	140
Respiratory Rate	breaths/min	8	20
Diastolic Blood Pressure	mmHg	50	90
Oxygen Saturation*	%	94	NA

* Oxygen saturation may be on supplemental oxygen

4.7 INTERIM ANALYSES

Other than the cumulative data review by the DMC for benefit/risk assessment, no formal efficacy interim analyses are planned at this time. The DMC, after reviewing unblinded data, may recommend that a formal efficacy interim analysis be performed. Furthermore, to adapt to information that may emerge during the course of this study, the Sponsor may choose to add a formal interim analysis at the recommendation of the DMC. Criteria of DMC requesting an interim analysis may be added to the DMC charter.

The interim analysis will be conducted by DMC. The decision to conduct the optional interim analysis, along with the rationale, timing, and statistical details for the analysis, will be documented in the Sponsor's trial master file prior to the conduct of the interim analysis. The DMC Charter will be updated to document potential recommendations the DMC can make as a result of the analysis (e.g., stop the study for positive efficacy, stop the study for futility).

5. REFERENCES

Cao B, Wang Y, Wen D, et al. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. N Engl J Med 2020. doi: 10.1056/NEJMoa2001282.

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

Lippi G, Plebani M. Laboratory abnormalities in patients with COVID-2019 infection. CCLM 2020. doi: 10.1515/cclm-2020-0198.

Fine, Jason P., and Robert J. Gray. A proportional hazards model for the subdistribution of a competing risk. Journal of the American statistical association 94.446 (1999): 496-509.

APPENDIX 1
PROTOCOL SYNOPSIS

PROTOCOL SYNOPSIS

TITLE:	**A PHASE II, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE SAFETY AND EFFICACY OF MSTT1041A OR UTTR1147A IN PATIENTS WITH SEVERE COVID-19 PNEUMONIA**
PROTOCOL NUMBER:	GA42469
VERSION NUMBER:	4
EUDRACT NUMBER:	2020-002713-17
IND NUMBER:	149507
NCT NUMBER:	NCT04386616
TEST PRODUCTS:	MSTT1041A (RO7187807) UTTR1147A (RO7021610)
PHASE:	II
INDICATION:	Severe COVID-19 pneumonia
SPONSOR:	Genentech, Inc.

Objectives and Endpoints

This study will evaluate the efficacy and safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with standard of care (SOC), for the treatment of severe coronavirus disease 2019 (COVID-19) pneumonia. Specific objectives and corresponding endpoints for the study are outlined below.

Efficacy Objectives

Primary Efficacy Objective

The primary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoint:

- *Time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first)*

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

Secondary Efficacy Objective

The secondary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoints:

- Time to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical status
- 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)
- Duration of supplemental oxygen
- Proportion of patients alive and free of respiratory failure (requiring non-invasive ventilation, high-flow oxygen, mechanical ventilation, or ECMO) at Day 28
- *Clinical status assessed using a 7-category ordinal scale at Days 14 and 28*
- Incidence of *invasive* mechanical ventilation *or* ECMO
- Ventilator-free days to Day 28
- Incidence of ICU stay
- Duration of ICU stay
- Time to clinical failure, defined as the time to death, mechanical ventilation, ICU admission, or withdrawal of care (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 rate at Days 14 *and* 28
- Time to clinical improvement, defined as a National Early Warning Score 2 of ≤ 2 maintained for 24 hours

Exploratory Efficacy Objective

The exploratory efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoints:

- Incidence of vasopressor use
- Duration of vasopressor use
- Incidence of starting hemodialysis
- Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral load on Day 15 and Day 28
- Proportion of patients with secondary bacterial infections

Safety Objective

The safety objective for this study is to evaluate the safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, 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)
- Change from baseline in targeted vital signs, targeted clinical laboratory test results, and targeted ECG parameters

Pharmacokinetic Objectives

The pharmacokinetic (PK) objective for this study is to characterize the MSTT1041A and UTTR1147A PK profiles on the basis of the following endpoints:

- Serum concentration of MSTT1041A at specified timepoints
- Serum concentration of UTTR1147A at specified timepoints

The exploratory PK objectives for this study are as follows:

- To evaluate potential relationships between drug exposure and the efficacy and safety of MSTT1041A and UTTR1147A on the basis of the following endpoints:
  - Relationship between serum concentration or PK parameters for MSTT1041A and efficacy and safety endpoints
  - Relationship between serum concentration or PK parameters for UTTR1147A and efficacy and safety endpoints
- To evaluate potential relationships between selected covariates and exposure to MSTT1041A or UTTR1147A on the basis of the following endpoint:
  - Relationship between selected covariates and serum concentration or PK parameters for MSTT1041A
  - Relationship between selected covariates and serum concentration or PK parameters for UTTR1147A

Immunogenicity Objectives

The immunogenicity objective for this study is to evaluate the immune response to UTTR1147A and MSTT1041A, individually, on the basis of the following endpoint:

- Prevalence of anti-drug antibodies (ADAs) at baseline and incidence of ADAs during the study

The exploratory immunogenicity objective for this study is to evaluate potential effects of ADAs on the basis of the following endpoint:

- Relationship between ADA status and efficacy, safety, or PK endpoints

Exploratory Biomarker Objective

The exploratory biomarker objective for this study is to identify and/or evaluate biomarkers that are predictive of response to MSTT1041A or UTTR1147A (i.e., predictive biomarkers), are early surrogates of efficacy, are associated with progression to a more severe disease state (i.e., prognostic biomarkers), can provide evidence of MSTT1041A or UTTR1147A activity (i.e., pharmacodynamic biomarkers), or can increase the knowledge and understanding of disease biology and drug safety, on the basis of the following endpoint:

- Relationship between biomarkers in blood and other fluid and efficacy, safety, PK, immunogenicity, or other biomarker endpoints

Study Design

Description of the Study

Overview of Study Design

This is a Phase II, randomized, double-blind, placebo-controlled, multicenter study to assess the efficacy and safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, in patients hospitalized with severe COVID-19 pneumonia. Approximately *390* adult hospitalized patients who have been diagnosed with COVID-19 pneumonia (defined by a positive polymerase chain reaction [PCR] and evidence of pneumonia by chest X-ray or computed tomography [CT] scan) are expected to be enrolled.

Patients will be randomized after screening at a 2:2:1:1 ratio to receive blinded treatment of either MSTT1041A, UTTR1147A, or their matching placebos. Study treatment will be given in combination with SOC as defined by the site, including, but not limited to, anti-virals, host-directed therapies, convalescent plasma, low-dose corticosteroids, and supportive care. Randomization will be stratified by need for invasive mechanical ventilation (yes vs. no) and region. Enrollment of patients with a need for invasive mechanical ventilation will be capped at approximately 25% of the overall study population.

Patients assigned to the MSTT1041A or UTTR1147A arm will receive one infusion of MSTT1041A 700 mg or UTTR1147A 90 μg/kg on Day 1, respectively, and patients assigned to the placebo arm will receive one infusion of matching placebo. A second dose of MSTT1041A 350 mg, UTTR1147A 90 μg/kg, or matching placebo will be given on Day 15 if the patient still remains hospitalized with a requirement for supplemental oxygen therapy. For patients who are being discharged or transferred to a different care facility prior to Day 60, a discharge visit should be performed. Subsequently, patients will be followed up remotely (via phone or video visit), and should return to the clinic for Day 28 and a study completion visit (Day 60) or early

discontinuation visit, if possible. Depending on patient findings during the phone/video visit, patients may need to be seen in person within 48 hours at the discretion of the investigator.

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.

Patients who are eligible to be re-screened will be required to repeat assessments as follows:

- Within the 2-day screening window: Repeat only the assessments that triggered screen failure.
- Outside the 2-day screening window: Repeat all assessments. The consent process does not need to be repeated if re-screening is completed within 7 days after completion of initial informed consent.

Note: Historic standard of care test results are acceptable for CT scan or chest X-ray if performed within 7 days prior to randomization, and for influenza and SARS-CoV-2 virology if performed within 14 days of randomization.

Number of Patients

Approximately *390* adult patients hospitalized with severe COVID-19 pneumonia will be enrolled in this study.

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Documented informed consent
- Age $\geq$18 years at time of signing Informed Consent Form
- Ability to comply with the study protocol
- Hospitalized with COVID-19 pneumonia confirmed per WHO criteria (including a positive PCR of any specimen; e.g., respiratory, blood, urine, stool, other bodily fluid) and evidenced by chest X-ray or CT scan
- For sites at an altitude $\leq$5000 feet: peripheral capillary oxygen saturation (SpO_2) $\leq$ 93% or partial pressure of oxygen/fraction of inspired oxygen $\leq$300 mmHg or requirement for supplemental oxygen to maintain SpO_2 $>$93%
- For sites at an altitude $>$5000 feet: requirement for supplemental oxygen to maintain SpO_2 at an acceptable level per local standard of care
- 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 95 days after the final dose of study drug.

  A woman is considered to be of childbearing potential if she is postmenarchal, has not reached a postmenopausal state ($\geq$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 95 days after the final dose of study drug 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:

- Pregnant or breastfeeding, or intending to become pregnant during the study or within 95 days after the final dose of study drug

  Women of childbearing potential must have a negative pregnancy test at screening.
- 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
- In the opinion of the investigator, progression to death is imminent and inevitable within the next 24 hours, irrespective of the provision of treatments
- Participating in another clinical drug trial
- Treatment with investigational therapy (other than for COVID-19) within 5 half-lives or 30 days (whichever is longer) prior to initiation of study drug
- Use of Janus kinase inhibitor within 30 days or 5 drug elimination half-lives (whichever is longer) prior to screening
- Have received high-dose systemic corticosteroids ($\geq$ 1 mg/kg/day methylprednisolone or equivalent) within 72 hours prior to Day 1
- Known HIV infection with CD4 $<$200 cells/μL or $<$14% of all lymphocytes
- ALT or AST $>10\times$upper limit of normal (ULN) detected at screening
- History of anaplastic large-cell lymphoma or mantle-cell lymphoma
- History of cancer within the previous 5 years unless it has been adequately treated and considered cured or remission-free in the investigator's judgment
- Clinical evidence of active or unstable cardiovascular disease (e.g., acute myocardial ischemia or decompensated heart failure) as assessed by the investigator
- Elevated cardiac troponin indicative of a recent cardiac event or myocarditis/pericarditis, as defined below:
  1. If high-sensitivity immunoassay is available locally: high-sensitivity troponin (hs-troponin) I or T $>$ULN (as per local standard for ULN), unless certain additional criteria are met, as outlined below:
     - If the local laboratory reports "indeterminate" or "intermediate" hs-troponin results: Patients with hs-troponin in the "intermediate" or "indeterminate" range (per local laboratory) may be enrolled if an echocardiogram shows normal left ventricular ejection fraction (as per local standard for normal, generally 50%–55%) without evidence of hypokinesis; if an echocardiogram cannot be obtained, clinical evaluation excluding myocarditis/pericarditis is acceptable.
     - If the local laboratory does not report "indeterminate" or "intermediate" hs-troponin results: Patients with hs-troponin $>$ULN to $<5\times$ULN may be enrolled if an echocardiogram shows normal left ventricular ejection fraction (as per local standard for normal, generally 50%–55%) without evidence of hypokinesis; if an

echocardiogram cannot be obtained, clinical evaluation excluding myocarditis/pericarditis is acceptable.

2. If high-sensitivity immunoassay is not available locally: conventional cardiac troponin I or T $>$ ULN, (based on local standard for ULN)

   – Patients with screen failure due to conventional troponin $>$ULN may be re-screened and enrolled if a repeat conventional troponin is $\leq$ULN and an echocardiogram shows normal left ventricular ejection fraction (as per local standard for normal, generally 50%–55%) without evidence of hypokinesis; if an echocardiogram cannot be obtained, clinical evaluation excluding myocarditis/pericarditis is acceptable.

- History or presence of an abnormal ECG that is clinically significant in the investigator's opinion, including complete left bundle branch block, second- or third-degree atrioventricular heart block, or evidence of prior myocardial infarction
- Sustained prolongation of QT interval corrected through use of Fridericia's formula (QTcF), defined as repeated demonstration of QTcF $>$480 ms (NCI CTCAE Grade 1)

  Patients with prolonged QTcF due to a reversible cause (e.g., electrolyte abnormalities) may be re-tested after the underlying cause has been corrected.

  For patients with a ventricular pacemaker, there should be appropriate correction for heart rate and pacing when determining baseline QTcF (as per Chakravarty et al. 2015); absolute QTcF values should not exceed 490 ms.
- History of ventricular dysrhythmias or risk factors for ventricular dysrhythmias such as structural heart disease (e.g., severe left ventricular systolic dysfunction, hypertrophic cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy), coronary heart disease (symptomatic or with ischemia demonstrated by diagnostic testing), or family history of sudden unexplained death or long QT syndrome
- History of moderate or severe allergic, anaphylactic, or anaphylactoid reactions or hypersensitivity to any component of study treatment

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 data point required for statistical analysis or safety 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 10 months.

Investigational Medicinal Products

Test Product (Investigational Drug)

- MSTT1041A 700 mg IV on Day 1, and a second dose of 350 mg IV on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time
- UTTR1147A 90 μg/kg IV on Day 1, and a second dose of 90 μg/kg IV on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time

Comparator

- Placebo for MSTT1041A via IV on Day 1, and a second dose on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time
- Placebo for UTTR1147A via IV on Day 1, and a second dose on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time

Statistical Methods

Primary Analysis

The primary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoint:

- *Time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first)*

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-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 ECMO or mechanical ventilation and additional organ support (e.g., vasopressors, renal replacement therapy)
7. Death

Time to recovery will be analyzed using the stratified log-rank test, adjusting for stratification factors. The Kaplan-Meier plot, median time to event, and their 95% CI and p-value from the stratified log-rank test will be presented. A Cox proportional hazards regression model will be used to estimate the hazard ratio comparing MSTT1041A or UTTR1147A with placebo, respectively, adjusting for stratification factors. Hazard ratios and 95% CIs will be presented. In addition, the p-value from unstratified log-rank test, unadjusted hazard ratio, and 95% CI will also be presented. Further details on the primary endpoint analysis will be included in the Data Analysis Plan.

Determination of Sample Size

A total of approximately *390* patients will be randomly allocated in a 2:2:1:1 ratio to receive MSTT1041A, UTTR1147A, or their matching placebos. The sample size provides approximately 80% power using a log-rank Chi-square test to detect a *7*-day difference between treatment groups in time to *recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first),* under the following assumptions: median time to improvement in the placebo group is *21* days, with 28 days follow-up, and using a one-sided 5% alpha. The minimal detectable difference is expected to be approximately *5.3 days.*

Optional Interim Analyses

Other than the cumulative data review by the Data Monitoring Committee (DMC) for benefit/risk assessment, no formal efficacy interim analyses are planned at this time. The DMC, after reviewing unblinded data, may recommend that a formal efficacy interim analysis be performed. Furthermore, to adapt to information that may emerge during the course of this study, the Sponsor may choose to add a formal interim analysis at the recommendation of the DMC.

The interim analysis will be conducted by DMC. The decision to conduct the optional interim analysis, along with the rationale, timing, and statistical details for the analysis, will be documented in the Sponsor's trial master file prior to the conduct of the interim analysis. The DMC Charter will be updated to document potential recommendations the DMC can make as a result of the analysis (e.g., stop the study for positive efficacy, stop the study for futility).

PROTOCOL SYNOPSIS

TITLE:	**A PHASE II, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE SAFETY AND EFFICACY OF MSTT1041A OR UTTR1147A IN PATIENTS WITH SEVERE COVID-19 PNEUMONIA**
PROTOCOL NUMBER:	GA42469
VERSION NUMBER:	4
EUDRACT NUMBER:	2020-002713-17
IND NUMBER:	149507
NCT NUMBER:	NCT04386616
TEST PRODUCTS:	MSTT1041A (RO7187807) UTTR1147A (RO7021610)
PHASE:	II
INDICATION:	Severe COVID-19 pneumonia
SPONSOR:	Genentech, Inc.

Objectives and Endpoints

This study will evaluate the efficacy and safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with standard of care (SOC), for the treatment of severe coronavirus disease 2019 (COVID-19) pneumonia. Specific objectives and corresponding endpoints for the study are outlined below.

Efficacy Objectives

Primary Efficacy Objective

The primary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoint:

- *Time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first)*

The ordinal scale categories are as follows:

8. 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)
9. Non–intensive care unit (ICU) hospital ward (or "ready for hospital ward") not requiring supplemental oxygen
10. Non-ICU hospital ward (or "ready for hospital ward") requiring supplemental oxygen
11. ICU or non-ICU hospital ward, requiring non-invasive ventilation or high-flow oxygen
12. ICU, requiring intubation and mechanical ventilation
13. ICU, requiring extracorporeal membrane oxygenation (ECMO) or mechanical ventilation and additional organ support (e.g., vasopressors, renal replacement therapy)
14. Death

Secondary Efficacy Objective

The secondary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoints:

- Time to improvement of at least 2 categories relative to baseline on a 7-category ordinal scale of clinical status
- 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)
- Duration of supplemental oxygen
- Proportion of patients alive and free of respiratory failure (requiring non-invasive ventilation, high-flow oxygen, mechanical ventilation, or ECMO) at Day 28
- *Clinical status assessed using a 7-category ordinal scale at Days 14 and 28*
- Incidence of *invasive* mechanical ventilation *or* ECMO
- Ventilator-free days to Day 28
- Incidence of ICU stay
- Duration of ICU stay
- Time to clinical failure, defined as the time to death, mechanical ventilation, ICU admission, or withdrawal of care (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 rate at Days 14 *and* 28
- Time to clinical improvement, defined as a National Early Warning Score 2 of ≤ 2 maintained for 24 hours

Exploratory Efficacy Objective

The exploratory efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoints:

- Incidence of vasopressor use
- Duration of vasopressor use
- Incidence of starting hemodialysis
- Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) viral load on Day 15 and Day 28
- Proportion of patients with secondary bacterial infections

Safety Objective

The safety objective for this study is to evaluate the safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, 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)
- Change from baseline in targeted vital signs, targeted clinical laboratory test results, and targeted ECG parameters

Pharmacokinetic Objectives

The pharmacokinetic (PK) objective for this study is to characterize the MSTT1041A and UTTR1147A PK profiles on the basis of the following endpoints:

- Serum concentration of MSTT1041A at specified timepoints
- Serum concentration of UTTR1147A at specified timepoints

The exploratory PK objectives for this study are as follows:

- To evaluate potential relationships between drug exposure and the efficacy and safety of MSTT1041A and UTTR1147A on the basis of the following endpoints:
  - Relationship between serum concentration or PK parameters for MSTT1041A and efficacy and safety endpoints

- Relationship between serum concentration or PK parameters for UTTR1147A and efficacy and safety endpoints

- To evaluate potential relationships between selected covariates and exposure to MSTT1041A or UTTR1147A on the basis of the following endpoint:
  - Relationship between selected covariates and serum concentration or PK parameters for MSTT1041A
  - Relationship between selected covariates and serum concentration or PK parameters for UTTR1147A

Immunogenicity Objectives

The immunogenicity objective for this study is to evaluate the immune response to UTTR1147A and MSTT1041A, individually, on the basis of the following endpoint:

- Prevalence of anti-drug antibodies (ADAs) at baseline and incidence of ADAs during the study

The exploratory immunogenicity objective for this study is to evaluate potential effects of ADAs on the basis of the following endpoint:

- Relationship between ADA status and efficacy, safety, or PK endpoints

Exploratory Biomarker Objective

The exploratory biomarker objective for this study is to identify and/or evaluate biomarkers that are predictive of response to MSTT1041A or UTTR1147A (i.e., predictive biomarkers), are early surrogates of efficacy, are associated with progression to a more severe disease state (i.e., prognostic biomarkers), can provide evidence of MSTT1041A or UTTR1147A activity (i.e., pharmacodynamic biomarkers), or can increase the knowledge and understanding of disease biology and drug safety, on the basis of the following endpoint:

- Relationship between biomarkers in blood and other fluid and efficacy, safety, PK, immunogenicity, or other biomarker endpoints

Study Design

Description of the Study

Overview of Study Design

This is a Phase II, randomized, double-blind, placebo-controlled, multicenter study to assess the efficacy and safety of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, in patients hospitalized with severe COVID-19 pneumonia. Approximately *390* adult hospitalized patients who have been diagnosed with COVID-19 pneumonia (defined by a positive polymerase chain reaction [PCR] and evidence of pneumonia by chest X-ray or computed tomography [CT] scan) are expected to be enrolled.

Patients will be randomized after screening at a 2:2:1:1 ratio to receive blinded treatment of either MSTT1041A, UTTR1147A, or their matching placebos. Study treatment will be given in combination with SOC as defined by the site, including, but not limited to, anti-virals, host-directed therapies, convalescent plasma, low-dose corticosteroids, and supportive care. Randomization will be stratified by need for invasive mechanical ventilation (yes vs. no) and region. Enrollment of patients with a need for invasive mechanical ventilation will be capped at approximately 25% of the overall study population.

Patients assigned to the MSTT1041A or UTTR1147A arm will receive one infusion of MSTT1041A 700 mg or UTTR1147A 90 μg/kg on Day 1, respectively, and patients assigned to the placebo arm will receive one infusion of matching placebo. A second dose of MSTT1041A 350 mg, UTTR1147A 90 μg/kg, or matching placebo will be given on Day 15 if the patient still remains hospitalized with a requirement for supplemental oxygen therapy. For patients who are being discharged or transferred to a different care facility prior to Day 60, a discharge visit should be performed. Subsequently, patients will be followed up remotely (via phone or video visit), and should return to the clinic for Day 28 and a study completion visit (Day 60) or early discontinuation visit, if possible. Depending on patient findings during the phone/video visit, patients may need to be seen in person within 48 hours at the discretion of the investigator.

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.

Patients who are eligible to be re-screened will be required to repeat assessments as follows:

- Within the 2-day screening window: Repeat only the assessments that triggered screen failure.
- Outside the 2-day screening window: Repeat all assessments. The consent process does not need to be repeated if re-screening is completed within 7 days after completion of initial informed consent.

Note: Historic standard of care test results are acceptable for CT scan or chest X-ray if performed within 7 days prior to randomization, and for influenza and SARS-CoV-2 virology if performed within 14 days of randomization.

Number of Patients

Approximately *390* adult patients hospitalized with severe COVID-19 pneumonia will be enrolled in this study.

Target Population

Inclusion Criteria

Patients must meet the following criteria for study entry:

- Documented informed consent
- Age ≥18 years at time of signing Informed Consent Form
- Ability to comply with the study protocol
- Hospitalized with COVID-19 pneumonia confirmed per WHO criteria (including a positive PCR of any specimen; e.g., respiratory, blood, urine, stool, other bodily fluid) and evidenced by chest X-ray or CT scan
- For sites at an altitude ≤5000 feet: peripheral capillary oxygen saturation (SpO_2) ≤ 93% or partial pressure of oxygen/fraction of inspired oxygen ≤300 mmHg or requirement for supplemental oxygen to maintain SpO_2 >93%
- For sites at an altitude >5000 feet: requirement for supplemental oxygen to maintain SpO_2 at an acceptable level per local standard of care
- 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 95 days after the final dose of study drug.

  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 95 days after the final dose of study drug 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:

- Pregnant or breastfeeding, or intending to become pregnant during the study or within 95 days after the final dose of study drug

  Women of childbearing potential must have a negative pregnancy test at screening.
- 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
- In the opinion of the investigator, progression to death is imminent and inevitable within the next 24 hours, irrespective of the provision of treatments
- Participating in another clinical drug trial
- Treatment with investigational therapy (other than for COVID-19) within 5 half-lives or 30 days (whichever is longer) prior to initiation of study drug
- Use of Janus kinase inhibitor within 30 days or 5 drug elimination half-lives (whichever is longer) prior to screening
- Have received high-dose systemic corticosteroids ($\geq$ 1 mg/kg/day methylprednisolone or equivalent) within 72 hours prior to Day 1
- Known HIV infection with CD4 $<$200 cells/μL or $<$14% of all lymphocytes
- ALT or AST $>10\times$upper limit of normal (ULN) detected at screening
- History of anaplastic large-cell lymphoma or mantle-cell lymphoma
- History of cancer within the previous 5 years unless it has been adequately treated and considered cured or remission-free in the investigator's judgment
- Clinical evidence of active or unstable cardiovascular disease (e.g., acute myocardial ischemia or decompensated heart failure) as assessed by the investigator
- Elevated cardiac troponin indicative of a recent cardiac event or myocarditis/pericarditis, as defined below:
  1. If high-sensitivity immunoassay is available locally: high-sensitivity troponin (hs-troponin) I or T $>$ULN (as per local standard for ULN), unless certain additional criteria are met, as outlined below:
     - If the local laboratory reports "indeterminate" or "intermediate" hs-troponin results: Patients with hs-troponin in the "intermediate" or "indeterminate" range (per local laboratory) may be enrolled if an echocardiogram shows normal left ventricular ejection fraction (as per local standard for normal, generally 50%–55%) without evidence of hypokinesis; if an echocardiogram cannot be obtained, clinical evaluation excluding myocarditis/pericarditis is acceptable.
     - If the local laboratory does not report "indeterminate" or "intermediate" hs-troponin results: Patients with hs-troponin $>$ULN to $<5\times$ULN may be enrolled if an echocardiogram shows normal left ventricular ejection fraction (as per local standard for normal, generally 50%–55%) without evidence of hypokinesis; if an

echocardiogram cannot be obtained, clinical evaluation excluding myocarditis/pericarditis is acceptable.

2. If high-sensitivity immunoassay is not available locally: conventional cardiac troponin I or T >ULN, (based on local standard for ULN)

   – Patients with screen failure due to conventional troponin >ULN may be re-screened and enrolled if a repeat conventional troponin is ≤ULN and an echocardiogram shows normal left ventricular ejection fraction (as per local standard for normal, generally 50%–55%) without evidence of hypokinesis; if an echocardiogram cannot be obtained, clinical evaluation excluding myocarditis/pericarditis is acceptable.

- History or presence of an abnormal ECG that is clinically significant in the investigator's opinion, including complete left bundle branch block, second- or third-degree atrioventricular heart block, or evidence of prior myocardial infarction
- Sustained prolongation of QT interval corrected through use of Fridericia's formula (QTcF), defined as repeated demonstration of QTcF >480 ms (NCI CTCAE Grade 1)

  Patients with prolonged QTcF due to a reversible cause (e.g., electrolyte abnormalities) may be re-tested after the underlying cause has been corrected.

  For patients with a ventricular pacemaker, there should be appropriate correction for heart rate and pacing when determining baseline QTcF (as per Chakravarty et al. 2015); absolute QTcF values should not exceed 490 ms.
- History of ventricular dysrhythmias or risk factors for ventricular dysrhythmias such as structural heart disease (e.g., severe left ventricular systolic dysfunction, hypertrophic cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy), coronary heart disease (symptomatic or with ischemia demonstrated by diagnostic testing), or family history of sudden unexplained death or long QT syndrome
- History of moderate or severe allergic, anaphylactic, or anaphylactoid reactions or hypersensitivity to any component of study treatment

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 data point required for statistical analysis or safety 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 10 months.

Investigational Medicinal Products

Test Product (Investigational Drug)

- MSTT1041A 700 mg IV on Day 1, and a second dose of 350 mg IV on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time
- UTTR1147A 90 μg/kg IV on Day 1, and a second dose of 90 μg/kg IV on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time

Comparator

- Placebo for MSTT1041A via IV on Day 1, and a second dose on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time
- Placebo for UTTR1147A via IV on Day 1, and a second dose on Day 15 for patients remaining hospitalized with a requirement for supplemental oxygen therapy at that time

Statistical Methods

Primary Analysis

The primary efficacy objective for this study is to evaluate the efficacy of MSTT1041A compared with placebo and of UTTR1147A compared with placebo, in combination with SOC, on the basis of the following endpoint:

- *Time to recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first)*

The ordinal scale categories are as follows:

8. 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)
9. Non-ICU hospital ward (or "ready for hospital ward") not requiring supplemental oxygen
10. Non-ICU hospital ward (or "ready for hospital ward") requiring supplemental oxygen
11. ICU or non-ICU hospital ward, requiring non-invasive ventilation or high-flow oxygen
12. ICU, requiring intubation and mechanical ventilation
13. ICU, requiring ECMO or mechanical ventilation and additional organ support (e.g., vasopressors, renal replacement therapy)
14. Death

Time to recovery will be analyzed using the stratified log-rank test, adjusting for stratification factors. The Kaplan-Meier plot, median time to event, and their 95% CI and p-value from the stratified log-rank test will be presented. A Cox proportional hazards regression model will be used to estimate the hazard ratio comparing MSTT1041A or UTTR1147A with placebo, respectively, adjusting for stratification factors. Hazard ratios and 95% CIs will be presented. In addition, the p-value from unstratified log-rank test, unadjusted hazard ratio, and 95% CI will also be presented. Further details on the primary endpoint analysis will be included in the Data Analysis Plan.

Determination of Sample Size

A total of approximately *390* patients will be randomly allocated in a 2:2:1:1 ratio to receive MSTT1041A, UTTR1147A, or their matching placebos. The sample size provides approximately 80% power using a log-rank Chi-square test to detect a *7*-day difference between treatment groups in time to *recovery, defined as time to score of 1 or 2 on the 7-category ordinal scale (whichever occurs first),* under the following assumptions: median time to improvement in the placebo group is *21* days, with 28 days follow-up, and using a one-sided 5% alpha. The minimal detectable difference is expected to be approximately *5.3 days.*

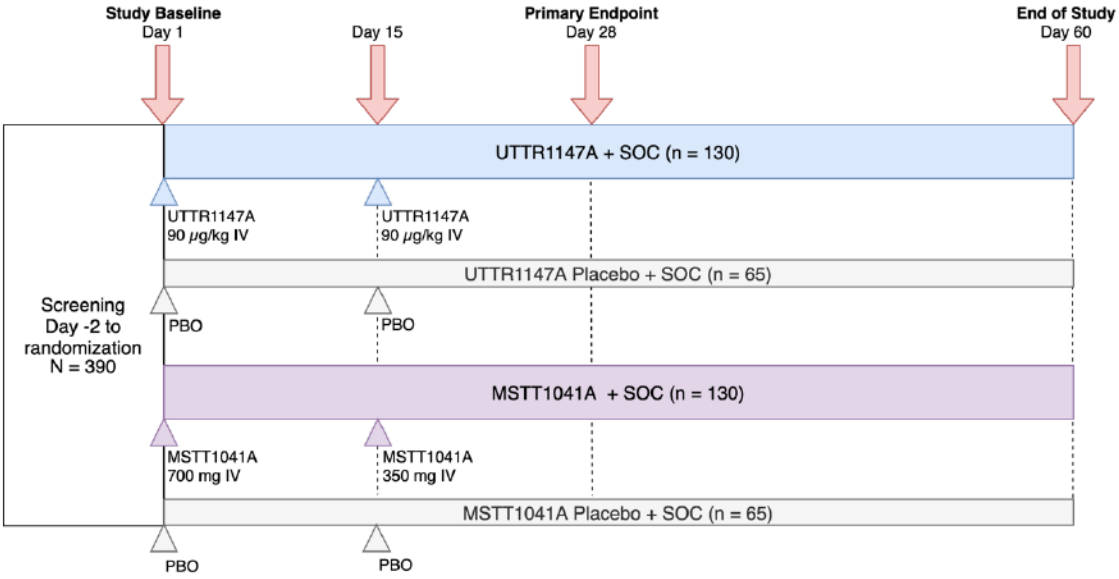
Optional Interim Analyses

Other than the cumulative data review by the Data Monitoring Committee (DMC) for benefit/risk assessment, no formal efficacy interim analyses are planned at this time. The DMC, after reviewing unblinded data, may recommend that a formal efficacy interim analysis be performed. Furthermore, to adapt to information that may emerge during the course of this study, the Sponsor may choose to add a formal interim analysis at the recommendation of the DMC.

The interim analysis will be conducted by DMC. The decision to conduct the optional interim analysis, along with the rationale, timing, and statistical details for the analysis, will be documented in the Sponsor's trial master file prior to the conduct of the interim analysis. The DMC Charter will be updated to document potential recommendations the DMC can make as a result of the analysis (e.g., stop the study for positive efficacy, stop the study for futility).

Appendix 2
Study Flowchart

Figure 1 Study Schema



PBO = placebo; SOC = standard of care.

Screening period will be up to 2 days.

A second dose of study drug will be given on Day 15 if the patient remains hospitalized with a requirement for supplemental oxygen therapy, unless patient meets the study drug discontinuation criteria (see Study Protocol for details).