Statistical Analysis Plan

A Multicenter, Adaptive, Randomized, Blinded Controlled Trial of the Safety and Efficacy of Investigational Therapeutics for Hospitalized Patients With COVID-19 (Trial H3: BRII-196/ BRII-198)

Version 3.1
23 February 2022

NCT05780424

Statistical Analysis Plan

Version 3.1

Therapeutics for Inpatients with COVID-19 (TICO)

A Multicenter, Adaptive, Randomized, Blinded Controlled Trial of the Safety and Efficacy of Investigational Therapeutics for Hospitalized Patients with COVID-19

Clinicaltrials.gov identifier: NCT04501978
EudraCT number: 2020-003278-37

Version	Date	Who	Comments
1.0	06 October 2020	BG	TICO ACTIV-3 INSIGHT 014 Protocol v1.0, 27 July 2020 Investigational agent: LY-Cov555
2.0	28 January 2021	BG	TICO ACTIV-3 INSIGHT 014 Protocol v2.0, 3 November 2020 and v3.0, 7 January 2021 Investigational agents: Vir-7831; BRII-196 and BRII-198; AZD7442
2.1	22 February 2021	BG	Added detail to win ratio analyses and risk score specifications
3.0	26 January 2022	BG	TICO ACTIV-3 INSIGHT 014 Protocol v4.0 (12 Mar 2021), v5.0 (9 Apr 2021), and v5.1 (17 Aug 2021) with letter of amendment #1 of 19 Aug 2021. Plans for biomarker analyses are included, based on the supplemental analysis plan dated April 14, 2021.
3.1	23 February 2022	BG	Added appendix for MP0420

Contents

1 Introduction

1.1 Objective of the Statistical Analysis Plan

The objective of this statistical analysis plan (SAP) is to provide a description of the general analytic strategy and the statistical methods that will be used to analyze the data for the TICO (Therapeutics for Inpatients with COVID-19) Phase III randomized, blinded, controlled, platform trial. This SAP applies to versions 3-5.1 of the TICO protocol. In version 2, two investigational agents are being studied, a SARS-CoV-2 neutralizing monoclonal antibody (nMAb) (Vir-7831) which is being developed by Vir Biotechnology (Vir) (San Francisco, CA) and GlaxoSmithKline (GSK) (Brentford, U.K.), and two nMAbs given sequentially (BRII-196 and BRII-198) which are being developed by Brii Biosciences (Durham, NC and Beijing). In protocol version 3, a third investigational agent is added, AZD7442, which is a combination of two nMAbs by AstraZeneca (Cambridge, U.K.); in version 4.0, MP0420, a DARPin® molecules was added; in version 5.0 (5.1), PF07304814, an antiviral protease inhibitor was added.

Participants are followed for death or re-hospitalization up to month 12 in version 2, and up to month 18 in versions 3, 4, and 5. The nMAbs and MP0420 are given by single infusion, or two sequential infusions for BRII-196 and BRII-198. PF07304814 requires a continuous infusion for 5 days.

The primary objective of the platform trial is to determine whether investigational agents, that are aimed at enhancing the host immune response to SARS-CoV-2 infection, or directly enhancing viral control in order to limit disease progression, are safe and superior to control (e.g., placebo) when given with standard of care (SOC) for the primary endpoint of time to sustained recovery evaluated up to 90 days of follow-up.

In the platform trial, several agents may be investigated in parallel, or staggered with overlapping times; investigational agents may be added or dropped. When more than one agent is being tested concurrently, where possible, participants will be randomly allocated across agents (as well as between the agent and its matched placebo), and the control group is pooled across the concurrently randomized, agent-specific matched placebo groups. Thus, each investigational agent and the corresponding pooled control group form their own randomized trial, and several agents may (at least partially) share their pooled control groups.

This SAP:

- Provides a short description of the study design (sections 1.2-1.4)
- Describes goals of the interim reviews by the independent DSMB and the planned format of the review meetings (section 2)
- Describes the planned data analyses presented in the reports to the DSMB (sections 3-13). General analysis principles are summarized in section 3, safety analyses are described in section 7, efficacy analyses in section 8, and interim monitoring guidelines in section 10.
- Describes data summaries to be provided regularly to study leadership to aid in monitoring trial conduct and data quality; these data summaries will be pooled across treatment groups, and will be restricted to enrolment, baseline data, and summaries of data completeness and study conduct.

The SAP for TICO will be updated by blinded study statisticians and clinical investigators following protocol amendments and prior to the unblinding the results for each investigational agent. For the latter, the blinded statisticians and clinical investigators will review data (pooled across treatment groups) on baseline characteristics, the number of participants who received a complete or partial infusion, and missing data due to withdrawal or loss to follow-up. A small group of the team will be unblinded to the pooled primary event rate to re-estimate sample size.

1.2 Description of the Study Design

This section is adapted from Section 1 of the TICO protocol versions 2.0 and 3.0.

Design

TICO is a master protocol to evaluate the safety and efficacy of multiple investigational agents aimed at modifying the host immune response to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, or directly enhancing viral control in order to limit disease progression.

Trials within this protocol will be adaptive, randomized, blinded and initially placebo-controlled. Participants will receive standard of care (SOC) treatment as part of this protocol. If an investigational agent shows superiority over placebo, SOC for the study of future investigational agents may be modified accordingly.

The protocol is for a phase III randomized, blinded, controlled platform trial that allows investigational agents to be added and dropped during the course of the study for efficient testing of new agents against control (i.e., placebo + SOC) within the same trial infrastructure. When more than one agent is being tested concurrently, participants will be randomly allocated across agents (as well as between the agent and its placebo). For analysis, placebo groups of concurrently randomized agents will be pooled; therefore, control groups may overlap for different agents.

Randomization will be stratified by study site pharmacy and disease severity. There are 2 disease severity strata, defined as below:

- **Disease severity stratum 1:** Absence of all of the following: stroke, meningitis encephalitis, myelitis, myocardial infarction, myocarditis, pericarditis, symptomatic congestive heart failure (NYHA class III or IV), arterial or deep venous thrombosis or pulmonary embolism, requirement for invasive mechanical ventilation, ECMO, mechanical circulatory support, vasopressor therapy, or new renal replacement therapy.

- **Disease severity stratum 2:** Presence of at least one of the excluded conditions or treatments in disease severity stratum 1.

The **primary endpoint** is the time from randomization to sustained recovery, defined as being discharged from the index hospitalization, followed by being alive and home for 14 consecutive days prior to Day 90. The definition of home will be operationalized as the level of residence or facility where the participant was residing prior to hospital admission leading to enrollment in this protocol.

An independent Data and Safety Monitoring Board (DSMB) will regularly review interim analyses that summarize safety and efficacy outcomes. For any agent, at the outset of the

trial, only participants in disease severity stratum 1 will be enrolled. This more restricted enrollment will continue until approximately 150 participants per study arm are enrolled and followed for 5 days. At this point, the DSMB will carry out a pre-specified assessment of futility, based on two 7-category ordinal outcomes (pulmonary and pulmonary+), assessed at Day 5. The pulmonary and pulmonary+ outcomes are described in Appendix A. Safety of the investigational agents will also be assessed.

For investigational agents passing this initial futility assessment, enrolment of patients in disease severity stratum 1 will continue, and it is planned to also expand enrollment, seamlessly and without any data unblinding, to include participants in disease severity stratum 2. The expansion to include more severely ill participants will be subject to recommendations by the FDA and the DSMB based on safety considerations.

After the initial futility assessment is passed, future interim analyses will be based on the primary endpoint of sustained recovery and will use pre-specified guidelines to determine early evidence of benefit, harm or futility for the investigational agent.

Primary Objective
The primary objective of this protocol is to determine whether investigational agents, aimed at enhancing the host immune response to SARS-CoV-2 infection, or directly enhancing viral control in order to limit disease progression, are safe and superior to control (e.g., placebo) when given with SOC for the primary endpoint of time to sustained recovery evaluated up to 90 days after randomization.

For all agents studied in TICO, the investigational agents/placebo have been given with study-supplied remdesivir unless contraindicated.

Duration
Participants will be followed for 18 months following randomization in versions 3-5 of the protocol. Primary and most secondary outcomes will be collected during the first 90 days of follow-up only. Follow-up beyond 90 days is planned because the half-lives of some agents indicate that potentially meaningful amounts may remain in the body after 90 days of follow-up. After 90 days through the end of follow-up, hospitalizations and deaths will be ascertained.

Sample size
This phase III trial is planned to provide 90% power to detect a 25% increase in the rate of sustained recovery for an investigational agent compared to placebo at the 0.025, 1-sided level of significance. This requires 843 primary events (i.e., participants who achieve sustained recovery). Randomization of 1,000 participants, equally allocated to each investigational agent and placebo, followed for 90 days is estimated to result in the required number of primary events. The event target may be achieved earlier if more than 1,000 participants are enrolled. Sample size will be evaluated periodically by study team members who are blinded to interim results on treatment difference and may be increased to maintain power for the hypothesized difference in sustained recovery between the investigational agent and placebo.

Population
The study population consists of inpatient adults (≥18 years) who have had COVID-19 symptoms ≤ 12 days. Initially, enrollment is restricted to disease severity stratum 1. After a pre-specified review by the DSMB for safety and futility, when approximately 150 participants

are enrolled per study arm, eligibility for randomization will be expanded to also include patients in disease severity stratum 2, subject to recommendations by the FDA and DSMB.

Stratification
Randomization is stratified by study site pharmacy; once enrolment is expanded to also include disease severity stratum 2, randomization will also be stratified by disease severity stratum.

Monitoring
An independent DSMB will review interim data on a regular basis for safety and efficacy. An initial futility assessment will be performed after the first 300 participants (150 in the active and 150 in the placebo arms) are enrolled and have Day 5 data; this initial assessment is based on two ordinal outcomes (pulmonary and pulmonary+ outcomes at Day 5). Afterwards, the DSMB will use pre-specified guidelines to identify agents with clear evidence of efficacy for the primary outcome and, if so, recommend unblinding of the trial results for that agent. Conversely, the DSMB may recommend discontinuation of an investigational agent if the risks are judged to outweigh the benefits or if futility assessments indicate that there is low probability that an investigational agent will achieve statistical significance for the primary endpoint of sustained recovery.

For an investigational agent, if the trial is stopped early or if the trial continues until the pre-specified number of primary endpoints is reached, further enrollment of the investigational agent will be terminated if applicable, and the trial data for the investigational agent will be unblinded and reported with data through 90 days of follow-up. Follow-up of all participants will continue through 18 months (12 months in version 2.0 of the protocol) using the data collection plan described in the master protocol.

Agent-specific Considerations

For individual investigational agents, the trial design may be modified, for example, the planned sample size for AZD7442 was increased from 1000 to 1500. Sample size was increased because data from the RECOVERY trial and the TICO trial of bamlanivimab indicated that the beneficial effect of antibody treatment could be based on the presence of neutralizing antibodies (nAbs) at entry, with nAb negative participants benefiting from antibody treatment and nAb positive participants showing no benefit or possible harm. In addition to increasing the sample size, the primary analysis for AZD7442 was modified, to test two hypotheses simultaneously while controlling the familywise type 1 error rate: treatment groups will be compared for time to sustained recovery in the full analysis population as well as in the subgroup of participants who were SARS-CoV-2 nAb seronegative at baseline using Holm's method for simultaneous tests. Such modifications are described in agent-specific appendices.

1.3 Randomization

The randomization is described in section 6.1 of the protocol.

Patients will be equally allocated to each investigational agent + SOC or to placebo + SOC. For example, for a study of a single investigational agent, participants will be randomized in a 1:1 ratio to the investigational agent or placebo. If a patient is eligible for two investigational agents and the placebo can be shared, the allocation will be 1:1:1 to investigational agent A, agent B, or placebo. Because the two investigational agents (A and B) may require different placebos

(for example, when infusion volumes differ), the 1:1:1 allocation ratio will be achieved through a two-step randomization procedure: in *step 1*, the participant is randomized 2:1 to "active" versus "placebo"; in *step 2*, the participant is randomized 1:1 to A versus B. With k agents, this can be viewed as an initial k:1 allocation to "active" versus "placebo", followed by a second, even allocation to one of the available agents (for example, if a participant was allocated to "placebo" in step 1, then the step 2 allocation will be 1:1 to "matched placebo for A" versus "matched placebo for B"). For the analysis, the concurrent agent-specific placebo groups will be pooled, resulting in a 1:1 allocation ratio for comparing each investigational agent versus the (pooled placebo) control group. If investigational agents are added or dropped, the allocation ratio to active versus placebo will be appropriately modified, and sample size will be recalculated as appropriate.

Randomization will be stratified by study site pharmacy (several clinical sites may share one pharmacy) and severity of disease at entry; the two disease severity strata are defined in section 1.2 of this SAP.

If more than one investigational agent is being compared with placebo and they have different contraindications, it is possible that a participant is eligible only for a subset of agents.

Comparisons will be of each investigational treatment against its control arm. The control arm consists of all participants who were "at risk" of being randomized to the investigational agent but were randomized to a control group instead. This concept is relevant when the randomization includes investigational agents with different eligibility criteria, when agents are introduced into the platform trial at different time points, or randomization to one of the agents is halted temporarily. Formal randomization includes agent-specific matched placebo groups, and the placebo groups will be pooled across agents, but only participants who 1) were eligible for the investigational agent under consideration, and 2) were randomized contemporaneously and at participating sites will be included in the control group for a given agent. At the time of randomization, for each participant, indicator variables will be set that record whether an agent was included in the randomization for that participant (e.g., indicator A=1, indicator B=1, indicator C=0 if the participant was eligible to be randomized to agents A and B, but not C). The pooled control group for agent A then consists of all participants who were randomized to (any) placebo, and for whom indicator A=1.

1.4 Sample Size Estimates

The planned sample size for each pairwise comparison is 1,000 participants (500 participants in each group). The sample size is sufficient to detect a recovery rate ratio (RRR) of 1.25 for time to sustained recovery with 90% power, using a one-sided test with a significance level of 0.025. The treatment groups are compared using Gray's test with ρ=0, the competing risks analogue of the log-rank test.

Sample size calculations are described in detail in Section 6.3 of the protocol.

Blinded sample size re-estimation will be carried out before enrollment is complete to determine whether the planned sample size of 1,000 participants followed for 90 days will yield the planned number of 843 primary events. The blinded sample size re-estimation does not involve unblinding of the treatment difference. It will be based on the pooled outcome data. Sample size may be increased to achieve the event target in the planned follow-up of 90 days. A

sample size increase may also be considered in order to achieve the event target before all participants are followed for 90 days.

When 300 participants (150 per study arm) are enrolled and have Day 5 outcome data, an early futility assessment is planned. This assessment will be based on two ordinal outcomes, denoted as “pulmonary” and the “pulmonary+”; both are ordered categorical outcomes with 7 categories, assessed on Day 5 (the two outcomes are described in Appendix A). Treatment groups will be compared using proportional odds models. With 300 participants, the early futility assessment is powered to detect a summary OR=1.60 with power of 95%, using a one-sided test with significance level of 0.30.[1] Given the two-outcome decision rules, an investigational agent with a summary OR=1.60 at Day 5 for both outcomes, would pass the futility assessment with a power between 93% and 98%, and a type I error between 0.21 and 0.39.

2 Interim DSMB Reviews: Goals and Format

Each investigational agent versus control will be reviewed as a separate clinical trial; separate data reports will be prepared for each investigational agent and the corresponding randomized (pooled) placebo group.

Goals of the interim reviews:

- Protect the safety of study participants.
- Advise on stopping or modifying the trial for efficacy, for patient safety in case of emerging data on harm, or for futility.
- After the first 300 participants are enrolled (150 participants per arm; all in disease severity stratum 1), advise on continuing the trial, and on expanding the study population to include more severely ill patients (disease severity stratum 2).
- Review the conduct of the trial
- If an investigational agent is stopped (due to efficacy, safety, or futility), the DSMB may be asked to advise on the timing of unblinding the data, in case the unblinding of the shared pooled placebo group may impact the integrity of the ongoing trial for another agent (section 14).

The DSMB will conduct frequent safety reviews. For an investigational agent with minimal pre-existing data, the first safety review will be conducted after 20-30 participants have been enrolled (10-15 per study arm) and Day 5 data are available, before increasing the pace of enrollment. Subsequent reviews will be timed according to the recommendations of the DSMB and study leadership. After the early futility review when 300 participants are enrolled (150 per study arm), further futility reviews would be expected to occur at approximately 50% and 75% information time (the number of observed sustained recoveries as a proportion of the targeted number of 843 events).

The DSMB may request interim reports that are focused on safety at any time.

Review meetings for each agent will typically consist of an Executive session (optional; closed), open session, closed session, and a second open session to give feedback to study leadership (optional). If several agents are reviewed at the same meeting, agents will be

reviewed consecutively, either with a sequence of open and closed sessions, or with one open session and one closed session (provided there are no unblinding conflicts).

Masking of treatment group labels in interim reports: In the open reports, any data reports will be pooled across the two treatment groups (the specific investigational agent and its pooled control group as described above). In the closed reports, treatment group labels will masked; for example as "Group A" versus "Group B". The treatment group labels will be consistent across all analyses and over subsequent reports. The DSMB will be unmasked to the treatment group labels.

Open report to the DSMB

The open reports for each investigational agent versus placebo comparison will contain:

- A synopsis of the trial design and current status of the platform trial
- Responses of the study team to DSMB requests
- A summary prepared by the study leadership
- Data summaries for enrolment, eligibility violations and protocol deviations, baseline characteristics
- Summary reports for data completeness and study conduct, pooled across treatment groups.
- Emerging external data, e.g., results of phase I or II trials on the investigational agent, will also be provided to the DSMB by the study leadership. This is usually included with the open report, but may be shared confidentially if needed.

All data summaries in the open report will be pooled across the investigational agent and placebo control. The open reports will be prepared by the blinded statisticians in cooperation with the unblinded statisticians. In addition to the DSMB, open reports will be provided to the study team, and posted on the website for access by study investigators.

While the study is ongoing, summaries by treatment group, and comparisons of the investigational agent versus placebo are restricted to the confidential closed report to the DSMB. Additionally, all summaries of follow-up data other than the data completeness and study conduct reports (pooled across the two treatment groups) will be restricted to the confidential closed report. For the **planned sample size re-estimations prior to completion**, the pooled number of primary events and the pooled event rate will be provided to the blinded study statisticians and study leadership. On a case-by-case basis, other pooled follow-up data may be provided if explicitly approved by the DSMB.

Closed report to the DSMB

All data summaries in the closed report will be by (masked) treatment group. The closed reports for a full review will contain:

- Specific data summaries requested by the DSMB or study leadership
- Data summaries in the open report, by treatment group (enrollment, baseline characteristics, eligibility violations)
- Data summaries to assess safety of the investigational treatment, described in sections 6 and 7. Data summaries for the primary "efficacy outcomes" and selected secondary outcomes will also be included in each report, because these data contain information about the risk/benefit profile of the investigational agent. Efficacy analyses are described in section 8.

- Subgroup analyses for sustained recovery and important safety outcomes will be conducted when the sample size is sufficiently large. Analyses are described in section 9.
- Data summaries on data completeness and study conduct, described in section 11
- Interim monitoring boundaries for efficacy or harm (section 10)
- Futility analyses (sections 10.2 and 10.4)
- Listings of grade 3 and 4 adverse events, serious adverse events (SAE), clinical organ failure and serious infections (PSEE), unanticipated problems (UP), suspected unexpected serious adverse reactions (SUSAR), and deaths.

Data reports will follow a similar format for all investigational agents. Each agent will have a small assigned team of unblinded statisticians, with 2-3 alternating teams when 2 or more agents are investigated in parallel. The unblinded statistician teams will cooperate in designing the master layout for the data reports and will serve as each other's backup when needed. The unblinded statisticians will be unblinded to several investigational agents in the platform trial; those for which they serve as primary statisticians, and those for which they serve as backup or advisory statisticians.

3 Analysis Principles

Each investigational agent versus control will be treated as a separate clinical trial; data reports will be for one "target" investigational agent and its corresponding randomized (pooled) control group. Investigational agents will not be directly compared against each other, unless explicitly stated in the agent-specific data analysis plan and agreed upon by all stakeholders. Therefore, in the event that several investigational agents are included in the platform trial in parallel, the pairwise comparisons of each agent versus control will **not** be adjusted for potential inflation of Type I error "due to multiple comparisons".

Comparisons will be of each investigational treatment against its (pooled) control arm.

Analysis populations:

- Comparisons for safety outcomes will be by modified intention-to-treat (mITT). The mITT analysis is restricted to participants who received a complete or partial infusion of the investigational agent/placebo; participants who did not receive any of the investigational agent/placebo are excluded.
- Comparisons for efficacy endpoints will be by intention-to-treat (ITT), unless otherwise stated. If analyses are by ITT, sensitivity analyses by mITT will be carried out for primary outcomes and key secondary outcomes.
- The analysis populations will be reconsidered prior to unblinding with consideration of the number of participants who were not infused, whether reasons for not infusing are independent of the treatment assignment, and the number of participants who have no follow-up data due to withdrawal of consent prior to infusion.

Pooled control group: As stated in section 1.3 above, the control arm for any investigational agent will be pooled across the agent-specific control groups for all agents that concurrently participated in the randomization. Specifically, the pooled control group for investigational agent A consists of all participants who might have been randomized to agent A but were randomized

to a placebo group instead. This concept is relevant when a participant is eligible to be randomized to more than one investigational agent, and agents were introduced into the platform trial at different time points or have different eligibility criteria.

In order to identify the pooled control group for each investigational agent correctly, the randomization application is setting indicator variables at the time of randomization for each participant that record whether an agent was included in the randomization (e.g., indicator A=1, indicator B=1, indicator C=0 if the participant was eligible to be randomized to agents A and B, but not C). The pooled control group for agent A then consists of all participants who were randomized to (any) placebo, and for whom indicator A=1.

Therefore, only participants who 1) were eligible for the investigational agent under consideration, 2) were randomized contemporaneously and at participating sites, and 3) were randomized to placebo will be included in the control group for a given agent.

Descriptive statistics will be reported overall and by randomized group. For categorical outcomes, the number and percent in each category will be reported; percentages will be of non-missing values, if data are not complete. Continuous variables will be summarized by median (interquartile range [IQR]) and/or mean (SD). Continuous variables may be categorized (e.g., age may be broken into categories to investigate the distribution across age groups).

Stratification: Tests comparing the investigational agent versus control for primary outcomes and key secondary outcomes will be stratified according to the planned randomization strata (disease severity and site pharmacy), provided participant numbers are sufficiently large. In this analysis plan, "stratification by disease severity" refers to the two disease severity randomization strata described in section 1.2. Initially, participants are enrolled only in disease severity stratum 1, until the investigational agent has passed the initial futility assessment (when 300 participants are enrolled for the pairwise comparison; 150 per study arm) and has been approved to expand enrolment to include participants from both strata 1 and 2. In this SAP, we use the notation "stratification by disease severity and site pharmacy" to denote the following:

- For analyses that include only participants in disease severity stratum 1, analyses will be stratified by site pharmacy.
- For analyses that include participants in both disease severity strata, analyses will be stratified by site pharmacy within each disease severity stratum; this means, the maximal possible number of strata is twice the number of site pharmacies. For analyses where stratification is implemented through addition of indicator variables for strata to the model, the strata would be defined through main effects for disease severity, for site pharmacy, and the interaction between disease severity and site pharmacy.

Because there are many site pharmacies, some strata may be small, particularly early in the trial. In order to avoid loss of power, any stratum that contains too few participants (less than 10-20 participants or events) should be pooled with other strata (of the same disease severity, and preferably within the same country or geographical region). Thus, several small strata may be pooled together, or pooled with a larger stratum.

For time-to-event analyses, if strata are too small for fitting separate baseline hazard functions, strata may be added as a categorical covariate to models instead. Whenever possible, however, analyses should be stratified by disease severity (with separate baseline hazard functions).

For **binary outcomes**, probabilities will be compared between the investigational agent and its control group using Cochran-Mantel-Haenszel tests (CMH) or logistic regression. If the numbers are sufficiently large, CMH tests will be stratified according to the planned randomization strata (disease severity and site pharmacy), as described above under "stratification". Odds ratios (OR) with 2-sided 95% confidence intervals (CI) will be estimated using logistic regression models.

For longitudinally measured binary outcomes, the treatment effect through follow-up will be estimated with 95% confidence intervals using generalized estimating equations (GEE) with a logit link function; the treatment effect is estimated via the interaction between the indicator for treatment group and the indicator for follow-up (versus baseline) visits. When there is more than one follow-up visit, "visit number" (day) may be included as categorical variable in the model, for variance reduction; alternatively, "time" may be included as a continuous variable.

Ordered categorical outcomes (pulmonary and pulmonary+) will be compared between treatment groups using proportional odds models, and the summary OR will be estimated with a 2-sided 95% CI.[2] Additionally, to aid the interpretation, the ordinal outcome will be dichotomized according to cumulative probabilities of the ordered categories, comparing treatment groups for proportions of participants in category 1, in the "best 2 categories", "best 3 categories", etc.; these comparisons will be performed using logistic regression (or stratified CMH tests).

For interim analyses of the ordinal outcomes, if one or more recently enrolled participants have died (i.e., their outcome status is known) but their current administrative follow-up has not yet reached the time point at which the treatment groups are to be compared (e.g., day 5), analyses will be restricted to participants with administrative follow-up greater or equal to the target time point.

Models will be adjusted for the baseline categories of the pulmonary+ outcome and for study pharmacy, by including the corresponding indicator variables in the model.

- If the number of observations is too small to adjust for both categorical covariates, preference will be given to the adjustment for the pulmonary+ category at baseline. Site pharmacy categories may be collapsed as described above under "stratification".
- For the initial futility analysis (after 150 participants per arm have Day 5 data), the adjustment for the pulmonary+ categories and pharmacy will be additive.
- For key analyses, unadjusted OR estimates will also be provided as sensitivity analyses.

The validity of the proportional odds assumption will be assessed by testing for heterogeneity in the log ORs (for the treatment effect) across the dichotomized cumulative ordered categories in the corresponding logistic regression model (partial proportional odds model, test for "unequal slopes").

- The primary sensitivity analysis testing the proportional odds assumption will compare the unadjusted proportional odds model for the treatment comparison (null model) versus a partial proportional odds model that allows for "unequal slopes" across the dichotomized cumulative categories (i.e., when testing the proportional odds assumption for the treatment comparison with respect to the pulmonary outcome on Day 5, the model will allow for heterogeneous ORs across the Day 5 pulmonary categories) as well as

across the stratification covariates (i.e., the baseline pulmonary+ categories and site pharmacy strata) (full partial proportional odds model).

Continuous outcomes will be compared between treatment groups using ANCOVA models for comparing means, if the ANCOVA model assumptions hold. If the distributions of the continuous outcomes are skewed, outcomes may be transformed, or compared between treatment groups using rank-based methods, such as the Wilcoxon test, or quantile (median) regression.

Comparisons between treatment groups for a continuous outcome will be adjusted for baseline values of the outcome, for the purpose of variance reduction, unless there are concerns over model stability with such an adjustment. For this purpose, the baseline value will be included as covariate in the model (e.g., ANCOVA, linear mixed models).

To estimate the treatment effect for longitudinally measured continuous outcomes, the outcome will usually be defined as "change from baseline" (difference at follow-up visit minus baseline value). The treatment effect through follow-up will then be estimated with 95% confidence intervals using generalized estimating equations (GEE) with an indicator for treatment group, or, in the case of Gaussian responses, the corresponding mixed effects models with random effects for participants. When there is more than one follow-up visit, "visit number" (day) may be included as categorical variable in the model, for variance reduction; alternatively, "time" may be included as continuous variable. Models will also be adjusted for the baseline values of the outcome variable.

Time-to-event outcomes will be summarized with Kaplan-Meier estimates for cumulative probabilities over time, and compared between treatment groups using log-rank tests or Cox proportional hazards models, or the corresponding competing risk analogues when death is a competing risk for the outcome. In particular, the primary endpoint of "time to sustained recovery" will be analyzed taking into account the competing risk of death. The following competing risk methods will be used:

- Aalen-Johansen estimator for the cumulative incidence function (analogue to the Kaplan-Meier estimate)[3]
- Gray's test with ρ=0 (analogue to the log-rank test)[4]
- Fine-Gray estimates and tests for the sub-distribution hazard ratio (analogue to the Cox proportional hazards model).[5,6]

The proportional hazards assumption will be tested by adding an interaction term for time by treatment group to the model. The cumulative proportions of participants who experienced the event will also be compared at given time points (specified in secondary objectives, e.g., at 28 days); in this case, the cumulative proportions will be estimated using Kaplan-Meier estimates or the competing risks analogue, and/or as proportion of participants who reached the time point (e.g., time since randomization $\geq$ 28 days).

The **administrative follow-up time** is defined as the minimum of (cut date minus randomization date) or the analysis time period. For example, the analysis time period for the primary endpoint of *sustained recovery* is 90 days, and the analysis time period for the important safety endpoint, the composite of *grade 3 and 4 events, SAEs, organ failure, serious*

infection or death, is 5 days or 28 days. The **administrative censoring date** is the earlier of the cut-date of the dataset, or randomization date plus analysis time period.

> **Comment:** The notion of "administrative censoring" is important in time-to-event analyses in the presence of competing risks. For example, the Fine-Gray method for estimating the sub-hazard ratio for sustained recovery can be approximated by using a Cox proportional hazards model where follow-up time for participants who died prior to achieving sustained recovery is not censored at death, but at the administrative censoring date.

Censoring for time-to-event analyses

For **interim** analyses, the type of censoring used will depend on the data collection schedule.

- If the reporting of the endpoint is data-driven (e.g., SAEs and deaths are reported as they occur), then follow-up is censored at the administrative censoring date, at the date of withdrawal, or loss to follow-up, whichever occurs earliest.
- If the date of the event is elicited retrospectively at fixed study visits spaced more than one week apart (e.g., "sustained recovery"), follow-up will be censored at the last day the endpoint status was ascertained.
- Sensitivity analyses will be provided for key analyses when the outcome status is uncertain.

For **final** analyses, follow-up will be censored on the last day the outcome status was ascertained.

Adverse events (AEs) will be classified by system organ class according to MedDRA®[1] (currently version 24.1 [September 2021] is used; when new versions are implemented, items are recoded). AEs will be graded according to the *DAIDS Table for Grading the Severity of Adult and Pediatric Adverse Events*, *Corrected Version 2.1 (July 2017)* (also referred to as the *DAIDS AE Grading Table*).[7] Cause of death will also be coded according to MedDRA®.

The number and percent of participants with grade 1-4 AEs will be summarized by day and grade, and by MedDRA® System Organ Class and grade. The percentage of participants with AEs will be compared between treatment groups according to grade cut-offs, e.g., "percent of participants with any AE", "percent of participants with grade 2 or higher AEs", etc., using CMH tests. The total number of events and median (IQR) of events per participant will also be summarized.

Additionally, the incidence of grade 3 and higher AEs will be summarized (number and percent of participants), and compared between treatment groups using time-to-event methods.

Significance level, two-sided tests: Unless noted otherwise, statistical tests and confidence intervals will be 2-sided, confidence intervals will have approximate 95% coverage probability, and test results with P-values $\leq$ 0.05 will be considered "significant". Percentages will be reported to at least one decimal place. P-values will be given to 2 significant figures.

Cut-date for interim reviews: Analysis data sets will be frozen (locked) several days (or weeks) prior to the review date, to allow the unblinded statisticians time to prepare a consistent

[1] The Medical Dictionary for Regulatory Activities terminology is the international medical terminology developed under the auspices of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). MedDRA® is a registered trademark of the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA)

report. The cut-date may be earlier than the date of the data freeze, to allow for lag time in the reporting of events. Early in the trial, the cut date and freeze date will be very close to the review date, to ensure timely safety reviews.

4 Enrolment and Eligibility

For the open report, the following enrolment and eligibility summaries will be provided:

- Enrolment over calendar time: plot by day or week, cumulative and increments.
- Enrolment by site pharmacy and by country: number (%)
- Eligibility: number (%) and reasons for eligibility violations

These summaries will be provided overall, and by disease severity randomization stratum.

For the closed report, enrolment and eligibility violations will be summarized by treatment group.

5 Baseline Characteristics

Baseline characteristics will be based on information collected on baseline and screening forms. For the open report, baseline characteristics will be summarized pooled across the two treatment groups (investigational agent and the "pooled" control group as described in section 2 above).

For the closed report, baseline characteristics will be summarized by treatment group.

The following baseline characteristics will be reported; unless noted otherwise, categorical variables will be summarized with numbers (%) in each category, and continuous variables will be summarized with median (IQR); in the open report, in addition, the mean (SD) and range may be provided. When enrolment has been completed, frequency distributions of these characteristics, including cut points for discretizing continuous measures, will be considered for defining subgroups in the final report.

- Demographics
  - Age: distribution in categories 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, ≥80 years; and summary as continuous variable
  - Sex at birth: number (%) male, female
  - Ethnic group: number (%) Asian, Black, Latino/Hispanic, White, other
  - Type of residence ("home")
  - Country of enrolment
- COVID-19 related characteristics
  - Duration of symptoms prior to enrolment
  - Use of remdesivir prior to enrolment
  - Pulmonary and pulmonary+ ordinal outcomes, number (%) in each category
  - NEWS: summary as continuous variable
  - Respiratory function scale (modified Borg dyspnea scale; continuous outcome)

  - Disease severity randomization stratum (for investigational agents that enrol in both strata), number (%) in each category
  - Receipt of SARS-CoV-2 vaccination, and type of vaccine (and if received as part of a blinded clinical trial, in which case the vaccine may be active or control)
- Other clinical characteristics
  - Concomitant treatments
  - Corticosteroid use will be summarized overall, and separately by oxygen requirement at baseline (no supplemental oxygen, < 4 L/min, conventional supplemental oxygen ≥ 4 L/min, high-flow oxygen or mechanical ventilation/ECMO)
  - History of chronic conditions (cardiovascular disease, diabetes, asthma, chronic obstructive pulmonary disease, hypertension, chronic kidney disease, hepatic impairment, cancer, or immunosuppressive disorder [HIV, and other than HIV])
  - Prior cerebrovascular event
  - Prior myocardial infarction (MI)
  - Requirement of continuous chronic supplemental oxygen
  - BMI (<30, 30-39.9, 40+)
  - Pregnancy (not applicable to protocol versions 2 and 3)
- Laboratory values: as continuous outcomes, and number (%) of grade 3 or 4 abnormalities according to the *DAIDS AE Grading Table.*

Some biomarkers will be measured centrally from stored samples, for example, hsCRP, IL-6, D-dimer, SARS-CoV-2 antibody, antigen, and viral RNA levels, and SARS-CoV-2 genome sequencing and variant identification. If these measures are available, they will be included in interim reports.

6 Administration of Study Treatment

These data are an important part of the safety review, with particular emphasis on infusion-related reactions and symptoms occurring during or within up to 2 hours after the infusion. These reactions and symptoms will be graded according to the DAIDS AE Grading Table.

The administration of study treatment is also an essential element of study conduct. Several summaries, pooled across treatment groups, will be included in the open report or provided to study leadership. Any summaries of adverse events or infusion-related reactions are restricted to the closed report.

For investigational agents administered as a one-time infusion, the following statistics will be used to summarize the infusion in each treatment group (active and control):

- Number and percentage of participants receiving complete infusion, partial infusion, infusion paused but resumed for complete infusion, or not infused (comparison by intention-to-treat).
- Number and percentage of participants with infusion-related reactions and symptoms (reported during the infusion or within 2 hours after the infusion), by grade. (Closed report only)
- Number and percentage of participants with an incident AE, SAE, UP or SUSAR on Day 0 during or after the infusion, overall and by oxygen requirement category at time of infusion (oxygen requirement at baseline, unless updated information is available; categories: no supplemental oxygen, < 4 L/min, conventional supplemental oxygen ≥ 4 L/min, high-flow oxygen or mechanical ventilation/ECMO). Types of AEs will be summarized by system organ class and by grade. (Closed report only)

- Number and percentage of participants who received:
  - Prior to infusion, medication to prevent infusion reactions, and type of medication
  - During or within 2 hours after infusion, medication to treat infusion reactions, and type of medication (Closed report only)
- Among participants infused, the day of infusion (same day as randomization, next day, > 1 day after randomization), and time between randomization and beginning of infusion (median hours, IQR).
- Among participants receiving full infusion, duration of infusion (median minutes, IQR).
- Time from vial puncture (beginning of preparation of the study agent by the pharmacist) to the end of the infusion, and number and percent of participants for whom the agent-specific time window was exceeded.
- Remdesivir:
  - Number and percent of participants who received (any) remdesivir, and number of days remdesivir was administered: median, IQR, distribution. (Closed report only)
  - Number and percent of participants who received remdesivir prior to the day of randomization, overall and by number of days prior
  - On the day of randomization: Number and percent of participants who received remdesivir prior to the investigational agent; after the investigational agent; no remdesivir.

Treatment groups will be compared by mITT (excluding participants who did not receive any investigational agent/placebo), unless specified otherwise. The treatment comparisons will be performed using the methods described in section 3 for binary and continuous outcomes (stratified CMH test for comparing percentages, and Wilcoxon rank-sum test [or quantile regression for comparing medians], respectively).

Selected summaries will also be provided separately for the two disease severity strata.

7 Safety Analyses

The planned timing of safety reviews is described in section 2. An overview of the safety data collection is provided in Appendix C.

Analysis cohort: Safety analyses will be carried out on participants who received a complete or partial infusion of the investigational agent (modified intention-to-treat [mITT]), unless otherwise stated.

A comprehensive safety review includes:
- Comparison of the treatment groups for the primary safety endpoint, its components, and analyses of secondary safety outcomes (described in this section)
- Analyses of infusion-related reactions and symptoms, described in section 6
- Evaluation of the “efficacy outcomes” (the pulmonary and pulmonary+ ordinal outcomes early in follow-up, and time to sustained recovery), which contain important safety information.

In addition to the full DSMB reviews, more frequent, shorter safety reports may be provided to the DSMB, for example, weekly safety reports early in the trial.

This section describes the primary safety outcome, and the analyses of AEs, SAEs, UPs, SUSARs, and deaths. Comparisons between treatment groups will be stratified by study pharmacy and by disease severity at study entry (as described in section 3 under "stratification").

In order to streamline the reporting of events, it was decided that certain protocol-specified exempt events (PSEE) are *not reported as SAEs*, unless they are considered related to the study treatment by the investigator. The events are listed in Appendix B; their composite is referred to as *clinical organ failure or serious infections*. While the *clinical organ failure or serious infection* events are serious events like SAEs, these events are reported not on the SAE eCRF unless they are considered related to the study treatment, but are reported as study endpoints on various other eCRFs.

The following safety and tolerability outcomes will be analyzed. If sample size permits, models will be stratified by disease severity and study site pharmacy, as described in section 3 under "stratification", unless noted otherwise.

- The **primary safety endpoint** is a composite of incident grade 3 or 4 clinical adverse events, SAEs, clinical organ failure or serious infections, or death through **Day 5**. The number and proportion of participants experiencing one of these events up through Day 5 will be tabulated, and treatment groups will be compared using a CMH test stratified by study site pharmacy and by disease severity at study entry.
    - Mortality will be analyzed as a key secondary outcome, see below.
    - The individual components of the composite outcome will be summarized.
    - *Sensitivity analyses*: For interim analyses, while the trial is still enrolling, treatment groups will also be compared for time to event through Day 5 using a log-rank test, stratified by site pharmacy and disease severity at study entry; the HR will be estimated with a 95% CI using a Cox proportional hazards model, and the cumulative proportion of participants with events over the first 5 days in each treatment group will be estimated using Kaplan-Meier curves.

  Comment: For investigational agents that are administered as a one-time infusion, such as nMAbs, the composite outcome at Day 5 is the primary measure to assess safety while the trial is still enrolling, as described in the protocol. This safety outcome and other safety outcomes through Day 28 will be considered along with the ordinal outcomes at Day 5 at interim data reviews by the DSMB to assess futility. When a trial is unblinded after the Day 90 follow-up is completed, however, treatment comparisons through Day 90 provide a more comprehensive evaluation of the investigational agent. For the publication of the main results for a specific agent, "time to death" and the composite of "death, SAE, clinical organ failure, or serious infections" through Day 90 would usually be presented as the main safety analyses, along with other safety outcomes. This decision will be agent-specific, and will be specified prior to the unblinding of the data for the agent.

- **All-cause mortality** through follow-up will be analyzed using time-to-event methods. Cumulative proportions of participants who died in each treatment group will be estimated using Kaplan-Meier estimates, and summarized in tables (proportion of participants who died by Days 5, 7, 14, 28, 60, 90, month 6, 12, and 18) and figures (Kaplan-Meier curves with pointwise 95% CIs). Treatment groups will be compared for time to death using log-rank tests, stratified by study site pharmacy and disease severity, and an overall HR will be

estimated with 95% CIs using stratified Cox proportional hazards models. Two other analyses will also be carried out to investigate the pattern of mortality differences, if any, between the treatment groups: 1) the proportional hazards assumption will be investigated; and 2) hazard ratios during the first 28 days of follow-up versus after 28 days will be estimated and compared.

- Cause of death will be MedDRA® coded and summarized by treatment group.

- The following composite endpoints will be analyzed using time-to-event methods (cumulative proportions of participants with events will be estimated using Kaplan-Meier curves with pointwise 95% CIs; treatment groups will be compared using log-rank tests; numbers and percent of participants with events will be summarized by treatment group, and overall HRs with 95% CI will be estimated using Cox proportional hazards models):
    - Composite of incident grade 3 or 4 clinical adverse events, SAEs, clinical organ failure or serious infection, or death through Day 28
        - Components of the composite endpoint will be also be summarized, overall and by system organ class. Proportions of participants who experienced any of these events by Day 28 will be compared using stratified CMH tests or logistic regression.
    - Composite of SAEs, clinical organ failure, serious infection, or death through Day 28 and Day 90
    - Composite of hospital re-admission or death through 18 months

- Treatment groups will be compared for the incidence of non-pulmonary events in the pulmonary+ ordinal outcome that are not part of the pulmonary outcome, through Day 5 and Day 7 (using time-to-event methods, with death as competing risk). These events are shown in red in Appendix A.

- AEs, SAEs, and UPs will be classified by MedDRA® system organ class. AEs will be graded for severity according to the *DAIDS AE Grading Table*. Grade 1-4 clinical AEs will be reported at baseline (Day 0 prior to infusion of the investigational agent), Day 0 after the infusion, Days 1-7, and on Days 14 and 28.

  The number and percent of participants with AEs will be summarized by day (Day 0 separately prior and after the infusion) and grade, and by system organ class and grade. Comparisons between treatment groups will be for the proportion of participants with AEs of a given grade or higher (i.e., any grade, grade 2+, grade 3+, grade 4). The treatment comparisons will be performed using stratified CMH tests or logistic regression.
    - For comparisons by day, the proportion of participants with any grade AEs will be compared for Days 0 (after infusion) through 7, and on Days 14 and 28.
    - For comparisons by grade, the proportion of participants who experienced any grade AEs (grade 2+ AEs, etc.) between Day 0 (after the infusion) through Day 7 will be compared.
    - For the comparison by system organ class, CMH tests will be performed if the number of participants with AEs is sufficiently large. System organ classes may be split up into MedDRA® preferred terms (PT) for system organ classes where the treatment difference is significant.

Other clinically meaningful AE groupings (beyond system organ class) may be developed by the study team, who are blinded to the treatment effect.

- In addition to any grade AEs through Day 7, grade 3 and 4 clinical AEs are being reported through Day 28. The number and percent of participants with incident grade 3 or 4 AEs through Day 28 will be summarized, overall and by system organ class, and compared between treatment groups using stratified CMH tests. (A grade 3 or 4 AE is considered "incident" if the event was not present at baseline or increased to grade 3 or 4 from grades 1 or 2.)

  To illustrate the time course, the incidence of *grade 3 or 4 AEs or death* through Day 28 will be summarized by treatment group using Kaplan-Meier estimates of the cumulative incidence functions (CIF).

- Infusion-related reactions and symptoms during infusion or within 2 hours after infusion of the investigational agent or placebo, and infusion cessation prior to completion will be tabulated and compared between treatment groups; analyses are described in section 6.

- Treatment groups will be compared for the proportion of participants who developed *organ failure or serious infections* through Day 28 and through Day 90, overall and by individual components, using stratified CMH tests. Individual components of this composite outcome will be tabulated.

- Treatment groups will be compared for incidence of organ failure, serious infections, or death through Day 28 and through Day 90, using Cox proportional hazards models; cumulative incidence functions will be estimated using Kaplan-Meier estimates.

- Treatment groups will be compared for incidence of a composite of fatal and non-fatal cardiovascular and thromboembolic events, a subset of the organ failure outcome (items 6b1, 6e2, 6e3, and 6f2 in Appendix B). Treatment groups will be compared using Cox proportional hazards models; cumulative proportions of participants with such events will be estimated using Kaplan-Meier estimates.

  Comment: Death for causes other than cardiovascular or thromboembolic events is a competing risk. As sensitivity analysis, the treatment groups will be compared using Aalen-Johansen estimates for the cumulative incidence functions and Fine-Gray models to estimate the sub-hazard ratio (investigational agent versus placebo) for the composite outcome. If results differ substantially from the Cox regression and Kaplan-Meier estimates, mortality for reasons other than CVD and thromboembolic events will be investigated in detail.

- Treatment groups will be compared for **new requirements of dialysis**, overall and in subgroups by eGFR level (<30 mL/min versus ≥30 mL/min at study entry), and subgroups by nAb status at study entry (nAb negative or positive).

  Motivation: In a joint analysis of data from the first three agents studied in TICO (bamlanivimab, sotrovimab, and BRII 196/198), the incidence of dialysis was elevated in the active treatment group compared with placebo (18 versus 2 events, HR=5.7 [95% CI: 1.2-24.6]). This safety signal was seen only among those who were nAb seronegative at study entry (16 events in the active arms versus 2 placebo, HR=5.0 [95% CI: 1.1 to 22.0]),

while only 2 events occurred among those who were nAb seropositive at study entry, both in the active arms. The mean time from enrollment to needing dialysis was 21 days.

Comment: Of the 20 participants with "incident dialysis", 2 had required dialysis prior to their COVID-19, and thus may be mis-classified for treatment-emergent dialysis. Both had eGFR < 10 mL/min at study entry and thus were at high risk for dialysis, and both had been allocated to active treatment.

Hypothesis: The incidence of dialysis is elevated with nMAb treatment compared with placebo, in particular among those with eGFR < 30 mL/min, and among those who are nAb seronegative at study entry.

- Treatment groups will be compared for mean changes in laboratory test values from baseline to Day 5, and for incidence of grade 3 and 4 laboratory abnormalities at Day 5 (new abnormality or increase in grade). Laboratory tests are conducted locally, and include serum creatinine, AST/SGOT or ALT/SGPT, WBC, hemoglobin, platelet counts, lymphocyte counts, and C-reactive protein. Statistical methods are described in section 3.

- Participants who are pregnant are not eligible for enrolment. For participants who become pregnant, pregnancy outcomes will be summarized.

- In addition to the safety outcomes specified in the platform protocol, other targeted safety outcomes for specific investigational agents may be specified in appendices to the protocol. Analyses will be specified in the corresponding agent-specific appendix to this SAP.

Listings of SAEs, clinical organ failure and serious infections, incident grade 3 and 4 AEs, UPs, SUSARs, and deaths (with cause of death) by treatment group will be provided at each DSMB meeting, with new events highlighted. The listings will include important baseline characteristics, such as age, sex, and disease severity (pulmonary outcome category) at study entry.

Further safety assessments may be considered.

Corticosteroid use will be monitored; concomitant medication use is collected at baseline and at Day 5.

- Corticosteroid use (any use at baseline or Day 5) will be summarized by treatment group and by oxygen requirement (worst category through Day 5: no supplemental oxygen, < 4 L/min, conventional supplemental oxygen $\geq$ 4 L/min, high-flow oxygen or mechanical ventilation/ECMO).
- Corticosteroid use on Day 5 will be summarized by treatment group and by oxygen requirement (worst category on Day 5, as above).

The impact of study arm on composite safety outcomes and their components at day 5, 28 and 90 will be assessed for subgroups defined by baseline characteristics, including SARS-CoV-2 antibody serostatus, antigen levels, demographics, duration of symptoms at enrollment, baseline classification of "home", clinical history and presentation (including disease severity stratum and pulmonary+ ordinal outcome at baseline). Tests for homogeneity of the treatment effect across subgroups will be carried out.

Outcomes and methods for subgroup analyses are described in detail in section 9.

8 Efficacy Analyses

Analysis cohort: Comparisons between each investigational agent and its concurrently randomized (pooled) controls will be by intention-to-treat (ITT) unless otherwise stated.

8.1 Primary Efficacy Endpoint and Primary Analysis

The **primary efficacy outcome** of the trial is "time from randomization to *sustained recovery* through Day 90"*. Sustained recovery* is defined as being discharged from the index hospitalization, followed by being alive and *home* for 14 consecutive days.

> **Comment:** The shortest possible time to sustained recovery is 14 days (this would require the patient to be discharged from the hospital on the day of randomization), and a patient would have to be discharged from the index hospitalization no later than Day 76 to achieve *sustained recovery* by Day 90.

Definition of *Home* for the primary endpoint:
According to the protocol, section 4.2, *Home* is defined as the level of residence or facility where the participant was residing prior to hospital admission leading to enrollment in this protocol.

Residence or facility groupings to define home are:

1) **Independent/community dwelling** with or without help, including house, apartment, undomiciled/homeless, shelter, or hotel
2) **Residential care facility** (e.g., assisted living facility, group home, other non-medical institutional setting)
3) **Other healthcare facility** (e.g., skilled nursing facility, acute rehab facility)
4) **Long-term acute care hospital** (hospital aimed at providing intensive, longer term acute care services, often for more than 28 days).

Lower (less intensive) level of residence or facility will also be considered as home. By definition, "home" cannot be a "short-term acute care" facility. Participants previously affiliated with a "long-term acute care" hospital recover when they return to the same or lower level of care.

Readmission from "home" (to a higher level of care) may occur and if this occurs within 14 days of the first discharge to "home", then the primary endpoint will not be reached until such time as the participant has been at home for 14 consecutive days.

Participants residing in a facility solely for public health or quarantine purposes will be considered as residing in the lowest level of required residence had these public health measures not been instated.

Primary analysis
The investigational agent will be compared to the (pooled) control group for *time to sustained recovery through Day 90* by intention-to-treat, using Gray's test with ρ=0.[4] The test will be

stratified by disease severity at entry and by site pharmacy (as described in section 3 under "stratification"). Gray's test compares the cumulative incidence functions for *sustained recovery* between the treatment groups, taking into account the "competing risk" of death in analyzing *sustained recovery*. Gray's test with ρ=0 is the competing-risks analogue of the log-rank test.

Comments:

1. Comparisons will be presented such that recovery rate ratios (RRR) >1 denote superiority of the investigational agent.
2. For AZD7442, the primary analysis is hierarchical: the overall comparison between treatment groups is followed by a comparison within the subgroup of participants who are nAb seronegative at study entry (described in Appendix E).

Analyses for the *sustained recovery* endpoint require methods that take into account the competing risk of death, as participants may die before ever achieving *sustained recovery*. The *sustained recovery* outcome requires knowledge of a participant's residence status for at least 14 days after arriving "home" (as defined above).

- The cumulative incidence functions for sustained recovery will be estimated by treatment group, using Aalen-Johansen estimators.[3] The estimates will be plotted over time, and tabulated at selected time points (days 15, 21, 28, 42, 60, 75, 90). The Aalen-Johansen estimator for a cumulative incidence function is the analogue of the Kaplan-Meier estimator in the presence of competing risks.

- The recovery rate ratio (RRR) for time to sustained recovery of the investigational agent versus control will be estimated, as a point estimate with a 95% CI, using the Fine-Gray model, stratified by disease severity at study entry and study site pharmacy.[5,6] The corresponding p-value for RRR=1 versus the two-sided alternative will be calculated. The Fine-Gray method is the competing risks equivalent of Cox proportional hazards models; the RRR compares the cumulative incidence rates of *sustained recovery* between the study arms and is a sub-distribution hazards ratio (sHR).

- To aid in the interpretation of the estimated treatment difference, the median days to sustained recovery (through Day 90) will be estimated for the investigational agent and the control group. Medians will be compared using the Wilcoxon rank sum test or quantile (median) regression. Participants who die at any time up to Day 90 will be assigned 91 days.

Censoring:

- Participants who are alive but have not experienced sustained recovery will be censored at the last date the endpoint status was ascertained (for interim analyses as well as the final analysis).
- For interim monitoring, two sensitivity analyses will be performed:
    1. Follow-up for time to sustained recovery will be censored *administratively* at the cut-date for the current report or Day 90, whichever comes first, with last known endpoint status carried forward. For participants who died, this type of censoring is integrated into Gray's test; using a log-rank test would require carrying forward the "not recovered" status for participants who died, up to the administrative censoring date.

2. Administrative censoring as described above will be applied, with the modification that participants who have been discharged from the hospital, were "home" at the latest date when residence was ascertained (but for < 14 days), and, if they have remained there, would have been at home for 14+ days by the cut-date, will be imputed as having experienced *sustained recovery* (achieved on day 14 at home).

Participants who withdrew consent or were lost to follow-up will be censored at the date of withdrawal or the last date the endpoint status was known, respectively.

In the first sensitivity analysis, the "not recovered" status is carried forward to the administrative censoring date; in the second analysis, "sustained recovery" is assumed at the earliest possible date. The first analysis potentially underestimates the rate of recovery, whereas the second analysis overestimates the recovery rate. In all analyses for time to sustained recovery, death is treated as competing risk.

Ascertainment of sustained recovery

The date of discharge from the index hospital will be recorded. Irrespective of the timing of the hospital discharge, there will be patient contact approximately every two weeks, on Days 14, 28, 42, 60, 75, and 90, either at a scheduled clinic visit or through phone contact. At these time points, a) vital status, and b) the location of the participant over time will be recorded, to assess whether the participant had been "at home" for 14 days. Therefore, the outcome status of sustained recovery will usually be ascertained within 3 weeks or less of the date the outcome was achieved.

- To illustrate the status of the primary endpoint, the recovery status of participants will be described over time with the following categories (at interim reviews):
  1. At home for 14+ days (reached the primary endpoint of sustained recovery)

  Did not reach sustained recovery, and:

  2. At home, < 14 days
  3. Discharged from the hospital, but not at home
  4. Hospitalized
  5. Dead
  6. Primary endpoint status unknown.

  The proportions of participants in each of the 6 categories will be summarized over time, by treatment group (stacked bar graphs and tables). In this analysis, both "sustained recovery" and "death" are absorbing states.

Assessment of model assumptions

- The trial was powered to detect an RRR of 1.25 with 90% power; this requires 843 sustained recoveries among the 1000 participants by Day 90. The rate of recoveries will be monitored, overall and within the two disease severity strata. Deviations of the observed distribution from the hypothesized distribution in the control arm will be monitored, and the impact on the power of the trial will be assessed. Prior to the completion of the trial for the investigational agent, sample size will be re-estimated by the blinded statisticians on the study team, based on the pooled rate of *sustained recovery*.

- The Fine-Gray model assumes that the sub-distribution rate ratio for *sustained recovery* is constant over time, similar to Cox proportional hazards models. The assumption of constant RRR will be tested by including an interaction effect between time and treatment indicator.

Sensitivity Analyses

- As sensitivity analyses, the primary comparison will be repeated after excluding participants who did not receive any of the investigational agent/placebo (modified intention-to-treat).
- Sensitivity analyses for the primary endpoint comparisons will include consideration of home oxygen above pre-morbid oxygen use (described in section 4.2.2 item 12 of the protocol). For these sensitivity analyses, *sustained recovery* will be re-defined as:
  a. "Discharged to home, alive at home without use of continuous supplemental oxygen for an uninterrupted 14 day period"
  b. "Discharged to home, alive at home for an uninterrupted 14 day period, and no supplemental oxygen use a the end of the 14 day period"
- If the RRR is not constant (test described under "assessment of model assumptions" above), as a sensitivity analysis, the RRR will be estimated within time periods, for example, Day 14-28, Day 29-60, Day 61-90.
- Additional sensitivity analyses are described under "censoring" above.

8.2 Key Secondary Outcomes

- Mortality is a key secondary outcome; analyses are described in section 7.

- To supplement the separate analyses of time to sustained recovery and time to death, the two endpoints will be analyzed jointly using the "win ratio" method for the composite outcome of time to recovery or death.[8] The win ratio will be calculated using the matched pairs method described in Pocock (2012).[8] Pairs will be formed by ranking the participants in each treatment group according to a risk score, described in section 13.2, and pairing the participants in groups A (here referring to the investigational drug) and B (referring to control) with equal ranks. Details are given below.

  - If both treatment groups have the same number of observations, the win ratio is calculated as follows:
    Step 1: Calculate the risk score for all participants, and order participants by the risk score in each treatment group. If needed, break ties at random. Each participant forms a "matched pair" with the participant of equal rank order in the other treatment group.
    Step 2: For each pair, determine whether the participant in group A wins, loses, or neither:
    a. *Compare pairs for time to death, for all pairs where one or both participants died.* If the participant in group A died, wins and losses are computed as follows:
       - If the matched participant in group B has longer follow-up, then A loses and B wins.
       - If the matched participant in group B has shorter follow-up and is alive at the censoring date, then neither group wins.

    Repeat for pairs where the participant in group B died.

b. *Compare remaining pairs for time to sustained recovery*.
   - If A achieved sustained recovery, and time to sustained recovery is longer for B, then A wins and B loses; vice versa for B.
   - If A achieved sustained recovery, and B was censored without reaching sustained recovery before A reached sustained recovery, then neither group wins; vice versa for B.
   - Otherwise, neither group wins.

Step 3: Calculate the win ratio as the number of wins in group A divided by the total number of pairs with a win or a loss in group A. Calculate the 95% CI for the win ratio and p-value as described in Pocock (2012).[8]

- If one treatment group has more participants than the other, select $|n_A - n_B|$ participants at random from the larger group and delete. Calculate the win ratio, 95% CI and p-value for the resulting matched pairs. Repeat the random selection of observations to delete 501 times; identify the matched pairs data set that corresponds to the median win ratio; the final values of the win ratio, 95% CI and p-value are those calculated from this data set.
- If both treatment groups have the same number of observations, but some ranked risk scores are tied within a treatment group, a similar process may be used to repeat the random breaking of ties, with the final win ratio chosen as the median over repeated random tie breaks.

With this approach, time to death is first used to determine the winning group (i.e., longer time to death), then time to sustained recovery is used to determine the winning group (i.e., shorter time to recovery): in this manner, the win ratio combines these conflicting outcomes into a composite while recognizing the importance of mortality.

8.3 Other Secondary Outcomes

The protocol defines a number of secondary endpoints in addition to the two key endpoints described in section 8.2 above. These analyses will be carried out for the final report. Selected secondary endpoints may also be analyzed for interim monitoring reports, to help evaluate the safety and efficacy of the investigational agent.

Below, the secondary outcomes from section 4.2.2 of the protocol are cited, with a short description of the analysis methods. For each outcome, the treatment group comparisons will be stratified by disease severity at study entry and by site pharmacy, as described in section 3 under “stratification”.

- Time to discharge for the initial hospitalization. Treatment groups will be compared using time-to-event methods that take into account the competing risk of death, similar to the analyses for time to sustained recovery described in section 8.1.
  - Hospital readmissions will be summarized using methods for recurrent events (i.e. those who are readmitted will re-enter the risk set).[9]
- Days alive outside of a short-term acute care hospital up to day 90. For this analysis, the “last-off” method will be used, i.e., days from the latest hospital discharge to day 90 will be counted. A person who dies within 90 days will be assigned a value of 0, consistent with the approach taken in trials of intensive care-based interventions. We will present

the median days by group and test the hypothesis of no difference between arms with a Wilcoxon rank sum test.

  - For interim analyses, only participants who have reached Day 90 (administrative follow-up for those who died) will be included, to avoid bias. Alternatively, a shorter time period may be used.

- Pulmonary+ and pulmonary ordinal outcomes on Days 1-7, and the pulmonary ordinal outcome on Days 14 and 28. The proportion of participants in each category of the pulmonary and pulmonary+ outcomes will be summarized over time (both outcomes at days 1-7, the pulmonary outcome also at Days 14 and 28); at each of those days, treatment groups will be compared using proportional odds models as described in section 3; the proportional odds models will be adjusted for the categories of the pulmonary+ outcome at baseline and for study site pharmacy. If participants in both disease severity strata are enrolled, the models will also be adjusted for the interaction between disease severity stratum and pharmacy.

  Additionally, the ordinal outcomes will be dichotomized ("category 1", "best 2 categories" through "best 5 categories"), and proportions will be compared between treatment groups at selected time points using logistic regression. For these analyses, the key dichotomized outcome considers the "best 2 categories", which is similar to the "recovery" outcome in the ACTT-1 trial.

- *Clinical organ failure or serious infections*, defined by development of any one or more of the clinical events listed in Appendix B, through Days 28 and 90. The development of *organ failure or serious infections* will be analyzed as a binary outcome, and the proportions of participants who developed *organ failure, serious infections or death* will be compared across arms using stratified CMH tests, overall and for individual components.

- A composite of death, clinical organ failure, or serious infection through Days 28 and 90 (see Appendix B). Treatment groups will be compared using standard time-to-event methods, since death is part of the outcome and not a competing risk. (Also described in section 7 as safety analysis).

- Outcomes assessed in other treatment trials of COVID-19 for hospitalized participants in order to facilitate cross-trial comparisons and overviews (e.g. 6- , 7-, and 8-category ordinal scales assessed at Days 1-7, 14 and 28; time to improvement in 1 or 2 categories of ordinal scale; time to best 3 categories of ordinal scale, and binary outcomes defined by improvement or worsening based on other ordinal outcomes). We will try to match the analyses in the other trials, to get results that can be compared. These analyses will not be performed for interim reports to the DSMB, unless requested.

- A composite of cardiovascular events (outcomes listed in items b1, e2 and e3 in Appendix B) and thromboembolic events (item f2). Time to event methods will be used that take into account the competing risk of death, e.g., Gray's test to compare treatment groups.

9 Subgroup Analyses

As stated in the protocol (section 11.2 in version 5.1), subgroup analyses for the primary efficacy outcome (time to sustained recovery) and for important safety outcomes will be performed. The goal is to determine whether and how the treatment effect (active versus control) differs qualitatively across various **subgroups defined at baseline**, and whether there are safety concerns in specific subgroups. Tests for heterogeneity of the treatment effect across subgroups will be carried out.

9.1 Outcomes, Subgroups, and Methods

The following outcomes will be considered for subgroup analyses:

- Time to sustained recovery (primary efficacy outcome)
- Time to the composite of *death, SAE, organ failure, or serious infections* through Day 90
- Time to death through Day 28, Day 90, and 18 months
- The composite of grade 3 and 4 events, SAEs, clinical organ failure, serious infections, or death through Day 5 and Day 28
- The Pulmonary and Pulmonary+ ordinal outcomes on Day 5
- Time to discharge from the index hospitalization, an outcome used in other COVID-19 trials.

To support the interpretation of the subgroup analyses for composite outcomes, components of the composites will also be analyzed.

Subgroup analyses will be performed by the following baseline factors:

- Disease severity (categories of the pulmonary+ outcome at study entry, randomization stratum). This subgroup analysis will be used at interim analyses after expansion of enrollment to assess if the treatment effect varies across the severity strata.
- Duration of symptoms prior to enrollment
- Age (18-49, 50-59, 60-69, 70-79, 80+)
- Biological sex
- Race/ethnicity
- Geographic location
- Level of residence (home) at the time COVID-19 symptoms developed
- Baseline pulmonary status (mutually exclusive subgroups: not on supplemental oxygen, supplemental oxygen < 4 L/min, supplemental oxygen $\geq$ 4 L/min, HFNC or NIV; invasive mechanical ventilation or ECMO)
- Body mass index (BMI)
- History of chronic conditions (cardiovascular disease, diabetes, asthma, chronic obstructive pulmonary disease, hypertension, chronic kidney disease, hepatic impairment, or cancer), and number of chronic conditions (none, 1, 2, 3 or more).
- Plasma antibody status. See section 9.2 below.
- Plasma antigen level. See section 9.2 below.
- SARS-CoV-2 viral RNA level based on midturbinate swab. See section 9.2 below.
- Biomarkers of inflammation and coagulation (including IL-6, hsCRP, and D-dimer)
- Concomitant medications, including subgroups formed by:

- Use of remdesivir prior to study entry (unless contraindicated, participants in both treatment groups were to receive remdesivir up to 10 days as background therapy)
- Use of corticosteroids (4 subgroup categories, formed by the combinations of supplemental oxygen use [yes/no] and corticosteroid use at study entry)

For the AZD7442 investigational agent, additional subgroup analyses will be carried out for the following subgroups defined at baseline; subgroups, hypotheses and analyses are described in Appendix E:

- SARS-CoV-2 vaccination status. Subgroups will be formed by vaccination status and by type of vaccine.
- Immunosuppressive state
- SARS-CoV-2 viral genome sequence and variant lineage.

As a **general rule**, subgroup analyses will be carried out using statistical models that are stratified by disease severity; for time-to-event analyses, stratified Fine-Gray or Cox proportional hazards models will be used, while for other outcomes, models will include the disease severity indicator as covariate. Global tests for heterogeneity of the treatment effect across subgroups will be carried out, by adding the interaction between the subgroup indicator and the treatment group indicator to a model that contains the treatment group indicator and subgroup variable as "main effects". In case the subgroup was formed by categorizing a continuous variable, the models used for the subgroup analyses will contain the (continuous) subgroup variable as covariate, and the interaction term in the expanded model will be formed between the subgroup indicator and the continuous variable.

Subgroup analyses for the primary endpoint of **time to sustained recovery** will use the Fine-Gray model (with death as competing risk), stratified by disease severity at study entry. Within each subgroup, the number and percent of recovered participants and median time to recovery will be calculated by treatment group, and sHRs with 95% CIs comparing the investigational agent versus control will be estimated. Global tests for heterogeneity of the treatment effect across subgroups will be carried out, by adding the interaction between the subgroup indicator and the treatment group indicator to the model that contains data across all subgroups. Median time to recovery will be estimated using Aalen-Johansen estimators.

Subgroup analyses for the **composite safety endpoint at Day 5** will use logistic regression, stratified by disease severity at study entry. Within each subgroup, the number and percent of participants with events will be calculated by treatment group, and ORs with 95% CIs comparing the investigational agent versus control will be estimated. Global tests for heterogeneity of the treatment effect across subgroups will be carried out, by testing for interactions between the treatment group and the subgroup variable in an expanded model, as described above. Additionally, subgroup analyses for the composite safety endpoint **through Day 5** will be conducted using Cox proportional hazards models, stratified by disease severity.

Subgroup analyses for **all-cause mortality** and **safety endpoints** that are analyzed using time-to-event methods (through Day 5, Day 28 and through Day 90) will use stratified Cox proportional hazards models, since death is part of the composite endpoints and not a competing risk. HRs will be estimated with 95% CIs for each subgroup, and global tests of heterogeneity of the treatment effect will be carried out, as described above. The proportional hazards models will be stratified by disease severity.

Subgroup analyses for the **pulmonary** and pulmonary+ ordinal outcomes on Day 5 will use proportional odds models, adjusted for the pulmonary or pulmonary+ category at study entry. In order to summarize the distributions across the categories of the ordinal outcome, the mean score (using 1-7 for best to worst) will be calculated by treatment group within each subgroup. For each subgroup, summary ORs with 95% CIs will be calculated. Global tests for heterogeneity of the treatment effect across subgroups will be carried out, as described above.

Subgroup analyses for **time to discharge from the index hospitalization** will use the Fine-Gray model (with death as competing risk), stratified by disease severity at study entry. Similar analyses as described above for time to sustained recovery will be carried out.

Additionally, subgroup analyses will be conducted for subgroups formed by a disease progression risk score at baseline. The construction of this risk score has been published and will be revisited as data on the sustained recovery endpoint accumulate for new investigational agents.[10]

Unless specified otherwise (e.g., the subgroups by nAb status at study entry for AZD7442 described in Appendix E), subgroup analyses will not be adjusted for multiple comparisons. Subgroup analyses will be interpreted with caution due to limited power and uncontrolled type I error.

9.2 Subgroups by Antibody Serostatus, Plasma Antigen, and Viral RNA Levels

An addendum to the SAP for version 1 of TICO, dated 14 April 2021, described analysis plans for antibody, antigen, viral RNA levels and biomarkers of inflammation and coagulation for monoclonal antibodies studied in TICO. This analysis plan was developed before the corresponding biomarker data were available. Results for the first three investigational agents, bamlanivimab, sotrovimab, and BRII-196/198, were published.[11] These analyses informed the current analysis plan for antibody, antigen, and viral RNA data.

This section motivates and describes subgroup analyses by baseline antibody serostatus, plasma antigen levels, and viral RNA levels. Assays are described in Appendix F. Briefly, plasma samples collected at study entry and at days 1, 3 and 5 were used to measure anti-spike receptor binding domain (RBD) neutralizing antibodies (nAbs) in a surrogate viral neutralization test, anti-nucleocapsid (anti-N) binding antibody levels, and semi-quantitative anti-Spike IgG levels. Qualitative and quantitative plasma SARS-CoV-2 N antigen was measured using a microbead-based immunoassay by Quanterix (Billerica, MA, USA). SARS-CoV-2 viral RNA levels were measured from a mid-turbinate nasal swab. SARS-CoV-2 genome sequencing and variant identification were also determined using Illumina whole genome sequencing of the mid-turbinate nasal swab at baseline.

Statistical models for the subgroup analyses depend on the investigated endpoint, and were described in section 9.1.

The two **key subgroup hypotheses** concern a differential treatment effect (active versus placebo) across subgroups by nAb serostatus for nMAb and antiviral treatments:

Hypothesis 1: Patients who are nAb seronegative (GenScript ELISA) at study entry will benefit more from the investigational agent (compared to placebo) than patients with higher endogenous antibody levels. Furthermore, those who are nAb seronegative AND have high antigen levels will benefit more from the investigational agent compared to placebo than other subgroups categorized by both antibody and antigen levels.

Hypothesis 2: Patients who are nAb seronegative (Genscript) AND have high levels of RNA from nasal turbinate swab will benefit more from the investigational agent compared to placebo than other subgroups categorized by both antibody and RNA levels.

To address these hypotheses, the following subgroups will be considered:

- Participants who are **nAb** negative at study entry, versus nAb positive. SARS-CoV-2 anti-spike neutralizing antibody levels by GenScript are expressed as percent binding inhibition. Specimens with levels < 30% are considered nAb negative (30% is the manufacturer's cutoff for positivity), specimens with levels ≥30% are considered nAb positive. In the TICO bamlanivimab trial, 50% of participants had endogenous nAbs (i.e., were nAb positive).
- Subgroups by **plasma antigen** level will be formed by using the median antigen level at study entry as a cut-point (< median versus ≥ median). In the TICO bamlanivimab trial, the median antigen level was approximately 1000 ng/L.
- Four subgroups formed by the bivariate combinations of nAb status (negative versus positive) and plasma antigen levels (< median versus ≥ median).
- Subgroups by **Viral RNA** levels by qualitative RT-PCR analysis on mid-turbinate nasal swabs will be formed using the median RNA levels (< median vs ≥ median). In the TICO bamlanivimab trial, the median was approximately 10,000 cp/mL.
- Four subgroups formed by the bivariate combinations of nAb status (negative versus positive) and viral RNA levels (< median versus ≥median).

Subgroup analyses will be performed for the **efficacy and safety outcomes** listed in section 9.1. The key efficacy outcome for the subgroup analyses will be time to *sustained recovery* through day 90, and key safety outcomes will be the composite of death, SAEs, organ failure and serious infections through day 90, mortality through day 90, and the composite of death, SAEs, organ failure, serious infections, and grade 3 or 4 AEs through day 28. Methods are described in section 9.1 above.

> **Comment:** In TICO, the neutralizing antibody status is determined at a central lab; while the baseline antibody status is being determined on an ongoing basis while the trial is enrolling, there is some delay due to batch shipping and processing time. Subgroup analyses by baseline nAb status will be included in interim reports to the DSMB when sufficient data are available for meaningful analyses.

Similar hypotheses as for the nAb subgroups will be assessed for subgroups formed by the additional two antibody assays; these analyses will be considered exploratory:

- Subgroups formed by **anti-N antibody status** (positive versus negative at study entry), and bivariate subgroups formed by the combinations of anti-N Ab status and plasma antigen or viral RNA levels. Anti-N antibody levels are measured using a SARS-CoV-2

antibody assay by **BioRad** (Hercules, CA, USA) measuring total (IgA, IgG, and IgM) anti-nucleoprotein. For the subgroup analyses, participants with anti-N levels <1 will be considered "anti-N Ab negative", those with levels ≥1 will be considered "anti-N Ab positive".

Comment: Results of the BioRad anti-N antibody measurement are defined in terms of "specimen ratios". According to the manufacturer, specimen ratios less than 0.8 are considered negative, those with a specimen ratio between 0.8 and 1.0 are considered equivocal, and those ≥ 1.0 are considered positive for the presence of anti-SARS-CoV-2 antibodies. For the subgroup analyses, participants with negative and equivocal anti-N Ab results will be combined and referred to as "anti-N negative".

- Subgroups formed by **quantitative anti-Spike IgG** (Quanterix) antibody status (positive versus negative), and bivariate subgroups formed by the combinations of anti-Spike IgG status and plasma antigen or viral RNA levels. Specimens with antibody levels < 770 ng/mL will be considered "anti-Spike IgG antibody negative", those with levels ≥ 770 ng/mL will be considered "anti-Spike IgG positive".

Motivation for the hypotheses:
In the TICO bamlanivimab trial, there was no difference in time to sustained recovery by day 90, RRR=0.99 (95% CI: 0.79 to 1.22), p=0.89. However, the treatment effect varied by nAb status at baseline; among patients who were nAbs negative and positive at entry, the RRRs were 1.24 (95% CI: 0.90-1.70) and 0.74 (95% CI: 0.54-1.00), respectively (p=0.02 for differential treatment effect). This means, there was a trend towards benefit with bamlanivimab among those who did not have endogenous nAbs at study entry, while time to sustained recovery was longer with bamlanivimab compared to placebo among those who already had endogenous nAbs. Moreover, among those who were nAb negative, the difference between bamlanivimab and placebo was more evident if plasma antigen or nasal-swab viral RNA were above the median levels at study entry. Results for subgroups by anti-N Ab status had a similar pattern, but less pronounced.

Motivation for safety analyses by nAb status:
In an analysis of the combined data for first three nMAbs investigated in TICO (bamlanivimab, sotrovimab, and BRII-196/198), the risk of the composite safety outcome of *death, SAE, organ failure, or serious infections through Day 90* was similar in the "active" versus placebo arms, but a safety signal was identified among participants with pre-existing endogenous neutralizing antibodies (nAbs). Specifically, the risk of the composite outcome at day 90 was 25.3% in the (pooled) active arms compared with 27.2% in the (pooled) placebo groups, HR=0.95 (95% CI: 0.72 to 1.25). The treatment effect on the composite safety outcome at day 90 differed by nAb serostatus at baseline, however; in participants who were nAb negative at baseline, HR=0.72 (95% CI: 0.50 to 1.02) (HR<1 favors the "active" group), and HR=1.42 (95% CI: 0.91 to 2.22) (favoring the placebo group) among those who were nAb positive at baseline, nominal p-value for interaction p=0.02. In the placebo group, the rate for the composite safety endpoint was higher among those who were nAb negative at baseline compared with those who were nAb positive (33% versus 20%, respectively); those who were nAb negative at baseline were sicker from COVID-19 at study entry and also had a lower chance of *sustained recovery*. Mortality followed a similar pattern as the composite outcome; there was no evidence for a difference between the active versus placebo groups in the overall analyses (pooled across the three nMAbs), HR=1.13 (95% CI:0.67 to 1.88), while the treatment effect on mortality differed

according to baseline nAb status, with HR=0.72 (95% CI: 0.39-1.33) (HR<1 favoring the active arm) among those who were nAb negative, compared to HR=3.64 (95% CI: 1.05 to 12.65) among those who were nAb positive.

10 Interim Monitoring Guidelines for the DSMB

Each investigational agent versus placebo comparison will be treated as a separate clinical trial; stopping boundaries will be derived to allow for multiple interim looks, but will not be additionally inflated to adjust for simultaneous analysis of multiple investigational agents, except when explicitly stated in the agent-specific protocol appendix and statistical analysis plan.

The DSMB will be asked to recommend early termination or modification only when there is clear and substantial evidence of a treatment difference (unless a trial is stopped for futility).

10.1 Early Assessment of Safety

For investigational agents with minimal pre-existing data, the pace of enrollment will be initially restricted and the DSMB will be asked to review safety data for the first 20 to 30 participants before increasing the pace of enrollment.

Subsequently, the DSMB will carry out regular reviews of safety data reports. These reports will include summaries of infusion–related events, grade 1-4 AEs, SAEs, organ failure, serious infections, and deaths, including the primary safety outcome at Day 5, and Day 28. Event listings for incident grade 3 and 4 AEs, SAEs, organ failure and serious infections, SUSARs, UPs and deaths will be provided (events that were reported since the previous review will be marked). Narratives will be provided for selected SAEs, SUSARs or UPs, particularly those judged related to study treatment. Analyses are described in section 7.

Monitoring boundary for harm:

- Until the first 300 participants are enrolled (150 per study arm) and the early futility analysis is conducted, the treatment groups will also be compared for the “pulmonary” and “pulmonary+” ordinal outcomes at Day 5, using a proportional odds model stratified by study pharmacy and pulmonary+ outcome category at baseline. A Haybittle-Peto boundary with 2.5 standard deviation (SD) for the first 50 participants enrolled and 2.0 SD afterwards will be used as a guideline for harm.

- After the study population is expanded to include disease severity stratum 2, these analyses are performed by disease severity stratum.

At the discretion of the DSMB, these safety reports will be prepared at a frequency they specify, for example, weekly. The DSMB may also request additional data summaries.

10.2 Early Assessment of Futility

10.2.1 Interim Monitoring Guidelines for Early Assessment of Futility

Early in the trial, enrolment is restricted to participants in disease severity stratum 1. When Day 5 data for the first (approximately) 300 participants (150 per study arm) are available, the DSMB will review interim data and use pre-specified guidelines for early evidence of sufficient activity of the investigational agent that justifies continuing enrolment for the agent and expanding eligibility criteria to include participants in disease severity stratum 2 as well as stratum 1. Because the risk-benefit trade-off may be more complex than can be captured in pre-specified guidelines, decisions to terminate an agent for futility will include a broad assessment of the available data, and may be postponed until more data is available.

The early futility monitoring uses two co-primary outcomes, denoted by “pulmonary” and “pulmonary+”, assessed on Day 5. Both are ordered categorical outcomes with 7 categories, described in section 4.1 of the protocol and in Appendix A to this SAP. The pulmonary outcome considers largely respiratory-related disease, similar to the ordinal outcome in the ACTT-1 trial.[12] The pulmonary+ outcome has the same categories for pulmonary complications (e.g., requirements for oxygen), and additionally includes extra-pulmonary outcomes such as thrombotic, myocardial, and cerebral complications of COVID-19.

Guidelines for the early futility assessment are as follows:

a. If the investigational agent is superior to placebo (i.e., $p \leq 0.3$ for a one-sided test) in both the pulmonary+ and pulmonary intermediate ordinal outcomes, then enrolment for the agent will expand to complete the trial.

b. If there is insufficient evidence for superiority versus control (i.e., one-sided $p>0.3$) in each of the two outcomes, then stop randomization.

c. If there is evidence (1-sided $p < 0.3$) for an association for one endpoint and not the other, then the agent may or may not advance depending on the risk/benefit profile emerging from the data at this early stage. If the effect estimate for both outcomes is on the side of benefit, the preference would be towards advancing the agent and expanding enrollment to include disease severity stratum 2.

   The DSMB will be asked to review whether the discordance is attributable to a positive or negative effect on extra-pulmonary organ dysfunction (the difference in the two ordinal scale categories, the conditions included in pulmonary+ but not in the pulmonary endpoint), and whether the same ordinal outcomes assessed on other days yield similar results, and weigh the risk/benefit profile. For example, if there is a significant positive effect on the pulmonary score and the lack of significant effect on the pulmonary+ score is driven by a lack of difference in the milder thrombotic symptoms in category 4 of the pulmonary+ scale (e.g. deep venous thrombosis) and there is no evidence of any raised risk of thrombosis overall, the agent will advance. Conversely, if the agent is superior to the control group with respect to the pulmonary outcome, but clearly inferior to the control group with respect to the pulmonary+ outcome or has a concerning safety profile, it will not advance.

Analyses of the primary efficacy endpoint, time to sustained recovery, will also be provided to the DSMB, as supporting information. These analyses are described in section 8.1.

If available, subgroup analyses by nAb status at study entry (nAb negative or positive) will also be provided.

10.2.2 Analyses of the Pulmonary and Pulmonary+ outcomes on Day 5

- Treatment groups will be compared by intention to treat.

- For each of the two ordinal outcomes, the number and percentage of participants in each of the categories on Day 5 will be tabulated, and the OR of the active versus control group will be estimated using a proportional odds model with indicators for the investigational agent group (active versus control) and for the categories of the ordinal pulmonary+ outcome at baseline (to adjust for baseline severity of illness).[2] The model will be stratified by site pharmacy.

  The summary tables will show the adjusted summary OR with 95% CI, estimated as described above, as primary analysis. In addition, the unadjusted summary OR with 95% CI will be shown as sensitivity analysis (estimated using a proportional odds model without adjustment for the pulmonary+ baseline category or site pharmacy). In the case that the adjusted OR differs substantially from the unadjusted OR, the reason for the deviation will be explored.

  Comments:
  - Results will be presented such that OR>1 favors the investigational agent, denoting higher odds of more favorable disease categories in the group randomized to investigational agent compared with control.

  - In order to avoid overestimating the proportion of participants who died, participants who died prior to Day 5 will only be included in the Day 5 summaries of the pulmonary and pulmonary+ outcomes if their time from randomization to cut-date is at least 5 days, and similarly for analyses on other days. Mortality is a key secondary endpoint and will be summarized cumulatively as an additional analysis.

- **For the initial futility analysis**, the tests comparing the investigational agent versus placebo will be performed using a (1-sided) type 1 error rate of 30%. This means, the investigational agent will be considered “superior” to the control with respect to the pulmonary (or pulmonary+) outcome, if the estimated summary OR is greater than 1, and the p-value $\leq$ 0.30.

  The summary reports will show the estimated summary OR with 95% CI, the signed Z-value for the test statistic comparing the treatment groups, and the one-sided p-value for superiority, calculated in the primary analysis (i.e., using the proportional odds model that is adjusted for the pulmonary+ category at baseline and stratified by site pharmacy, as described above). As sensitivity analysis, these values will also be calculated using the unadjusted proportional odds model and included in the summary report.

  Comment: At the recommendation of the FDA, for monoclonal antibodies studied, patients requiring high-flow oxygen or mechanical ventilation (invasive or non-invasive) are not eligible for enrolment unless the investigational agent passes the initial futility assessment. At that time, the FDA will be consulted. This was done for AZD7442 after it passed the initial futility

assessment and the FDA allowed expansion of enrollment. This initial enrollment restriction does not apply to investigational agents other than monoclonal antibodies.

If eligibility is restricted to disease severity stratum 1, adjusting the treatment comparison for the pulmonary+ outcome at baseline is identical to adjusting for the following categories defined by oxygen requirement:

- No supplemental oxygen (pulmonary+ category 2)
- Supplemental oxygen < 4 L/min (or < 4 L/min above premorbid requirements) (pulmonary+ category 3)
- Supplemental oxygen ≥ 4 L/min (or ≥ 4 L/min above premorbid requirements, but not high-flow oxygen) (pulmonary+ category 4)

- To supplement the overall summary odds ratios for the 7-category outcomes, each dichotomized definition of improvement that can be formulated from the components of the ordinal outcomes will be considered separately; for example, treatment groups will be compared for the proportions of participants in category 1 on Day 5, proportions in categories 1 or 2 ("best two categories"), in categories 1-3, etc. Proportions will be tabulated, and odds ratios for active versus control groups will be estimated with 2-sided 95% CIs using logistic regression models. These analyses need to be interpreted with caution, because they are not adjusted for inflation of type I error due to multiple comparisons.

- Subgroup analyses will be carried out for the Pulmonary and Pulmonary+ outcomes on Day 5, to supplement the early futility analyses. The goal is to determine whether the treatment effect differs across subgroups, and to aid the DSMB in considerations on whether there are safety concerns in specific subgroups. Principles for subgroup analyses are described in section 9.1; here, subgroup analyses are based on the proportional odds models. In particular, heterogeneity of the treatment effect across the baseline pulmonary+ categories will be assessed.

- After an investigational agent has passed initial futility assessment and enrollment has been expanded to include participants in disease severity stratum 2, treatment comparisons for the pulmonary and pulmonary+ outcomes on Day 5 continue, and will be performed separately for each of the two disease severity strata, to assess safety for the more severely ill participants in stratum 2.

Missing data: Unknown outcome status for the pulmonary or pulmonary+ outcomes on Day 5:

The following items describe how missing data will be treated for the primary analyses of the pulmonary or pulmonary+ outcomes on Day 5. As needed, these methods may be also applied to analyses at other time points (e.g., Day 7).

- **Interim analyses**:
  - Only participants with Day 5 data for the pulmonary outcome will be included for the Day 5 comparisons. The number and proportion of participants with unknown outcome status will be summarized.
    Comment: If the cut date is less than 10 days before the data freeze date, Day 5 data for the ordinal outcomes are considered "missing" only for participants with at least 10 days of administrative follow-up.

- **Final analyses** after completion of the trial:
  - If Day 5 data are missing for a substantial proportion of participants (e.g., more than 5% of the mITT cohort), multiple imputation will be used to impute missing Day 5 data for the pulmonary and pulmonary+ outcomes. For the imputation, the following baseline covariates will be considered in addition to the indicator for treatment group: age, sex, country, duration of symptoms prior to enrollment, status of the ordinal pulmonary (or pulmonary+) outcome, and presence of comorbidities. Ten rounds of imputation will be used to estimate the summary odds ratio.
  - The number and proportion of participants with missing data will be reported.

Sensitivity analyses

- As sensitivity analyses, the treatment groups will be compared by **modified intention-to-treat (mITT)** after excluding participants who did not receive any of the assigned investigational agent (active or control). This mITT analysis will be provided at important decision points, e.g., when the test statistic approaches the monitoring boundary, and for the final analyses after completion of the trial.
- Treatment groups will be compared for the pulmonary and pulmonary+ outcomes on Days 1-7, to monitor the consistency of the treatment effect over time.

Assessment of model assumptions

- For the pulmonary and pulmonary+ outcomes at Day 5, the proportionality assumption of the odds ratio will be assessed (by including the interaction between the treatment group indicator and indicators for the Day 5 cumulative ordinal categories in the model, as well as the interactions between the treatment group indicator and the indicators for the strata by baseline pulmonary+ categories and site pharmacy; this corresponds to testing for separate slopes using a partial proportional odds model, see section 3 under "ordered categories"). If there is evidence for non-proportionality, the summary odds ratio in the proportional odds model will still be used to quantify the treatment effect, and the analyses of the dichotomized ordinal outcome categories will be used to help interpret the treatment effect.

- The sample size of 300 (150 per study arm) is sufficient to detect a summary OR of 1.60 for the comparison of the investigational agent versus control for each of the two ordinal outcomes with 95% power. The power of the tests depends on the hypothesized OR and the hypothesized distribution in the control group used for the sample size calculations. At the time of the early futility review, the deviation of the observed distribution from the hypothesized distribution in the control arm will be assessed, and the impact on the power of the trial will be estimated.

10.3 Interim Monitoring Guidelines for the Primary Endpoint

This section describes the interim monitoring guidelines that apply after the investigational agent has passed the initial futility assessment (after approximately 150 participants per study arm are enrolled).

As a guideline, asymmetric boundaries will be provided to monitor the primary endpoint (time to sustained recovery) for overwhelming benefit or for harm. The trial of an investigational agent should be stopped for efficacy only if there is clear and convincing evidence of superiority

of the agent versus the pooled control group with respect to the primary outcome, time to sustained recovery. For monitoring superiority, the Lan-DeMets spending function analogue of the O'Brien-Fleming boundaries will be used, with a 1-sided 0.025 level of significance over multiple looks. For computing the Lan-DeMets boundary, the information fraction at each interim analysis will be the observed total number of sustained recoveries divided by the planned number of sustained recoveries (N=843).

The monitoring boundary for harm is asymmetric, requiring less evidence to stop for harm than for superiority; a Haybittle-Peto boundary with 2.5 SD for the first 50 participants enrolled and 2.0 SD afterwards will used as a guideline for harm. With this approach, less evidence will be required for crossing a boundary for harm than for benefit.

At each full interim review after the first 169 participants have achieved sustained recovery (20% information time), the following will be provided:

- Signed square root of the value of the test statistic for Gray's test with ρ=0, ("Z-value") comparing the investigational agent versus the control group for the primary endpoint through Day 90, plotted over information time, and the asymmetric monitoring boundaries: the O'Brien-Fleming boundary with Lan-DeMets α-spending function for superiority (one-sided test with α=0.025), and the asymmetric, Haybittle-Peto boundary for harm described above.

    - **Comment:** Test statistics for the primary treatment comparison will be coded such that the value of the test statistic > 0 favors the investigational agent. Thus, in case of harm, the Haybittle-Peto boundary with 2 SD of the normalized test statistic, is crossed if the Z-value for the test statistic is below -2, irrespective of information time.

  In addition to the current value of the test statistic, the corresponding values of the test statistic at the previous reviews will be plotted over information time, (1) as presented at the previous DSMB meetings, and (2) re-calculated with current data (using the cut-dates of the previous reports).

- History of the estimated rate ratios for time to sustained recovery with 95% CIs and p-values (by Fine-Gray's method), and normalized test statistic values and p-values for Gray's test at previous DSMB reviews, as presented, and recalculated with the current data (using the cut-date of the previous reports). The latter provides information on the influence of a possible time lag in the ascertainment of sustained recovery.

10.4 Interim Monitoring for Futility

After investigational agents have passed the initial futility assessment (based on the pulmonary and pulmonary+ outcomes at Day 5, assessed for the first 300 participants), further futility analyses will be based on the primary outcome of *time to sustained recovery*.. The aim of these analyses will be to consider whether an investigational agent should be discontinued due to a low probability of achieving statistical significance for the primary endpoint of sustained recovery at the completion of the 90 day follow-up.

Conditional power calculations for time to sustained recovery will be presented under a range of scenarios. In the primary futility analysis, it will be assumed that the treatment effect for the

future, as yet unobserved follow-up will be as hypothesized in the study design (RRR=1.25). As secondary analysis, the treatment effect for future follow-up will be assumed to be similar to the observed effect. Additional scenarios may be provided. Typical futility guidelines recommend stopping a trial when conditional power (assuming the originally hypothesized treatment effect for the future follow-up) is below 10%-15%.[13]

As a guideline, futility will first be assessed when 50% of the planned number of sustained recoveries have occurred, and a value of 15% will be suggested as a threshold for the conditional power. An additional assessment will take place at 75% of the events. Conditional power will be computed using Gray's test with ρ=0, the competing risk analogue of the log-rank test.[14]

Decisions to terminate an agent for futility will include a broad assessment of the risk/benefit trade-off in addition to these guidelines.

11 Data Completeness and Study Conduct

According to the protocol, the pulmonary and pulmonary+ outcomes will be assessed on days 0-7; the decision rules for at the initial futility assessment are based on these outcomes on Day 5. The pulmonary outcome will also be assessed on Days 14 and 28. The primary outcome, "time to sustained recovery", will be assessed through Day 90. Clinical data will be collected on Days 0-7, 14, 28, 60 and 90; mortality and re-hospitalizations will be assessed through 18 months. After hospital discharge, in-person visits are scheduled on Days 1, 3, 5, 28, and 90, when blood is collected (plasma and serum); other visits may be conducted by phone (Days 7, 14, 42, 60, and 75). The data collection schedule is included in Appendix D of this SAP.

Data completeness and study conduct reports will be provided by treatment group (for the closed report) and pooled across treatment groups (for the open report). Data summaries for the infusion of the investigational agent on Day 0 are described in Section 6; several of those reports are also relevant for monitoring study conduct and will be included in the open report or provided to study leadership, pooled across treatment groups.

The following data summaries will be provided to assess data completeness and study conduct:

- Number and percent of participants with protocol deviations, and type of protocol deviation
- Expected and observed number (% of expected) of participants who completed visits on Days 1-7, 14, 28, 42, 60, 75, and 90.
- Expected and observed number (% of expected) of participants with known outcome status for the pulmonary and the pulmonary+ outcomes on Day 5.
- Ascertainment of the primary outcome: Expected and observed number (% of expected) of participants with known status of "time to sustained recovery" at days 28, 60, and 90. To ascertain "sustained recovery", several elements are required: vital status; the status of hospitalization; if discharged, the status of the residence ("home" versus other).
- Expected and observed number (% of expected) of participants with known vital status at days 5, 14, 28, 60 and 90, and at months 6, 12, and 18.

- Number and percent of participants who withdrew consent or were lost to follow-up (no contact and unknown vital status for 45+ days).
- If substantial numbers of participants are lost to follow-up (e.g., more than 10% of participants), Kaplan-Meier estimates for the cumulative proportion of participants who are lost to follow-up over time, by treatment group, will be provided (closed report only).
- Listing of participants who withdrew consent, including dates of randomization, pulmonary+ category at baseline, receipt of study treatment, date of withdrawal, and reason of withdrawal.
- Length of follow-up: Median, IQR, range
- Collection of specimens: Expected and observed number (% of expected) of participants with specimens collected as specified by the protocol, by visit.
- Expected and observed numbers of participants with local laboratory data at baseline and on Day 5.

A visit counts as "expected" if the visit window has closed or the data have been received.

12 SARS-CoV-2 Antibody, Antigen, and Viral RNA Levels

SARS-CoV-2 neutralizing antibody (nAb) levels, anti-N antibody levels, and plasma antigen levels are measured in stored plasma specimens collected at baseline (Day 0), Day 1, Day 3, and Day 5. SARS-CoV-2 RNA levels are determined from mid-turbinate nasal swabs, collected at baseline. Assays are described in Appendix F.

The statistical analyses described in this section are based on the analysis plan for antibody and antigen data for bamlanivimab (LY-CoV555), the *Addendum to Statistical Analysis Plan (Version 1.0) Therapeutics for Inpatients with COVID-19 (TICO) ACTIV-3 INSIGHT 014*, dated 14 April 2021. The current update to the SAP is further informed by the completed analyses of antibody, antigen, and viral RNA data for bamlanivimab, sotrovimab, and BRII 196/198, the first three investigational agents in TICO.[11]

Antibody, antigen, and viral RNA levels are determined centrally, and may not be available at interim analyses. If data are available, key analyses will be included in interim reports.

12.1 Associations of Baseline Levels with Clinical Outcomes

For each treatment group, associations of baseline antibody status (negative versus positive), levels of plasma antigen, and levels of viral RNA at study entry, with time to *sustained recovery* through day 90 (primary efficacy outcome) and with important safety outcomes will be assessed.

12.2 Plasma Antigen Levels Through Day 5

Plasma antigen levels over time (Baseline, Days 1, 3, and 5) will be described using the following summary statistics, by treatment group, for all participants with follow-up data, and separately for participants who were nAb negative versus nAb positive at study entry:

- Number and percent of participants with antigen levels < 3 ng/L ("antigen negative"), and number and percent of participants with antigen levels < median antigen levels at

study entry (the cut-point used for subgroup analyses).
- Side-by-side box plots of antigen levels. Antigen levels < 3 ng/L (the lower limit of quantification) will be imputed as 2.9, and antigen levels will plotted on the log_{10}-scale.
- Geometric mean concentrations (GMC) of the antigen levels

Hypothesis 1: Over the first 5 days, antigen levels will decline faster in those receiving the investigational agent compared to those receiving placebo.

Hypothesis 2: The treatment effect (investigational agent versus placebo) on the antigen levels is stronger among participants who are nAb negative at study entry compared with those who are nAb positive at study entry.

To address these hypotheses, treatment groups will be compared for the percentage of participants who are antigen negative (i.e., antigen levels < 3 ng/L) at Days 1, 3, and 5, overall (Hypothesis 1) and within subgroups (Hypothesis 2).

- At each time point, the treatment difference will be estimated using a logistic regression model with the log_{10}-transformed baseline antigen level and disease severity indicator as covariates, and ORs with 95% CIs and p-values will be cited.
- In addition to the treatment comparisons at each time point (Days 1, 3, and 5), longitudinal GEE models will be fitted for the proportion of participants who are antigen negative over time, and slopes over time will be compared between treatment groups. The GEE models will be adjusted for baseline levels and disease severity.
- In addition to estimating the treatment differences within each subgroup, heterogeneity of the treatment effect across subgroups by baseline nAb status (Hypothesis 2) will be assessed by testing for an interaction between the treatment group indicator and the subgroup indicator for baseline nAb status in an expanded GEE model.
- Unadjusted analysis will also be performed, as sensitivity analysis.

Comment: Within the full study population, treatment comparison will be adjusted for nAb status at study entry by including the nAb status and disease severity indicators as covariates. Within nAb subgroups, comparisons will be adjusted for disease severity.

As secondary analyses, treatment groups will also be compared for the percentage of participants with antigen levels < median antigen level at study entry, as well as for the geometric mean concentration (GMC) of antigen levels.

In order to compare treatment groups for geometric mean concentration (GMC) of antigen levels, antigen levels < 3 ng/L will be imputed as 2.9 ng/L, and the treatment difference will be estimated as the GMC ratio for the investigational agent versus placebo. For this analysis, plasma antigen levels will be log_{10}-transformed, mean differences will be estimated with 95% CIs in ANCOVA models that include the treatment indicator as fixed effect and the baseline log_{10} antigen level and disease severity indicator as covariates, and the treatment difference will be back-transformed to the original scale to estimate the GMC ratio. These analyses will be performed in the full study population, as well as within subgroups by nAb status. Heterogeneity of the treatment effect across subgroups will be assessed by testing for an interaction between treatment group and subgroup indicator in expanded ANCOVA models for the log_{10}-transformed antigen levels.

12.3 Antibody Levels Through Day 5

Levels of nAb and anti-N antibodies at baseline, Days 1, 3, and 5 will be described using the following summary statistics, by treatment group, for all participants with follow-up data, and separately within subgroups by antibody status (positive versus negative) at study entry. To investigate nAb levels over time, subgroups will be formed by nAb status. To investigate anti-N antibody levels over time, subgroups will be formed by anti-N antibody status. Similar analyses will be performed for anti-Spike IgG antibody levels.

- Number and percent of participants with nAb percent binding ratio $\geq$ 30% ("nAb positive")
- Number and percent of participants with anti-N Ab specimen ratio $\geq$ 1 ("anti-N antibody positive")
- Number and percent of participants with anti-Spike IgG antibody level $\geq$ 770 ng/mL ("anti-Spike IgG positive")
- Side-by-side box plots of nAb, anti-N, and anti-Spike IgG antibody levels over time
- Median concentrations of the nAb, anti-N, and anti-Spike IgG antibody levels

Comment: The manufacturer of the anti-N antibody assay (BioRad) considers samples with specimen ratios < 0.8 as "antibody negative", and samples with values between 0.8 and 1 as "equivocal". For our analyses, we refer to all samples with specimen ratio < 1 as "anti-N negative".

Treatment groups will be compared for the following outcomes, overall and within subgroups by antibody status at study entry:

- The percentage of participants who are nAb positive (i.e., percent binding ratio $\geq$ 30%) at Days 1, 3, and 5, overall and within subgroups by nAb status at study entry.
- The percentage of participants who are anti-N antibody positive (i.e., specimen ratio $\geq$ 1) at Days 1, 3, and 5, overall and within subgroups by anti-N antibody status at study entry.
- The percentage of participants who are anti-Spike IgG antibody positive (i.e., antibody level $\geq$ 770 ng/mL) at study entry at Days 1, 3, and 5, overall and within subgroups by anti-spike IgG antibody levels at study entry.

Statistical methods are similar to those described for comparing antigen levels over time.

In addition, at Days 1, 3, and 5, treatment groups will be compared for mean antibody levels (nAb, anti-N, and anti-Spike IgG levels), and the mean difference (investigational agent minus placebo) will be estimated with 95% CIs, overall and in subgroups by antibody status, using ANCOVA models that include the treatment indicator as fixed effect, and the baseline antibody level (continuous covariate) and disease severity indicator as covariates. Treatment groups will also be compared for mean change over time using longitudinal mixed effects models, comparing slopes of mean antibody levels over time between treatment groups. Heterogeneity of the treatment effect across subgroups will be assessed by testing for an interaction between treatment group and subgroup indicator in expanded longitudinal mixed effects models.

Comment: The analyses of mean levels of nAb and anti-N antibody levels are exploratory, as the assays are qualitative. In contrast, the anti-Spike IgG antibody assay is semi-quantitative, and for that assay percentiles will be considered in subgroup analyses as well as classification of results as antibody negative/positive.

13 Exploratory Analyses

13.1 Associations Between the Pulmonary and Pulmonary+ Outcomes and Time to Sustained Recovery

At the early futility assessment when 300 participants (150 per study arm) are enrolled and have Day 5 data, estimated treatment differences in the pulmonary and pulmonary+ ordinal outcomes on Day 5 are used to identify promising investigational agents to continue to be studied with the clinical outcome of "time to sustained recovery". In exploratory analyses, we will investigate whether the early pulmonary and pulmonary+ outcomes on Day 5 are an adequate predictor for time to sustained recovery, to re-evaluate our initial futility decision rule.

- Associations between the pulmonary and pulmonary+ ordinal outcomes at Day 5 with time to sustained recovery will be estimated using Cox proportional hazards models, pooled across treatment groups. To illustrate these associations, median time to sustained recovery will be estimated by category of the ordinal outcomes on Day 5, using Aalen-Johansen estimates of the cumulative incidence function.

Ideally, the goal would be to evaluate the extent to which treatment differences in the pulmonary outcomes on Day 5 predict treatment differences in time to sustained recovery through Day 90. A detailed analysis plan will be developed at a later time.

13.2 Disease Progression Risk Score

A disease progression risk score, calculated at baseline, will be used to form subgroups of participants with low or high predicted risk for subgroup analyses for safety and efficacy outcomes, and to pair participants for the win ratio analyses described in section 8.2.

The risk score for a participant is defined as the estimated *probability that the Pulmonary outcome on Day 5 is in one of the categories 5, 6 or 7* (5=non-invasive ventilation or high-flow oxygen, 6=invasive ventilation, 7=death). The probability is estimated using a logistic regression model for the corresponding binary outcome (Pulmonary categories $\geq$5 vs <5 on Day 5) with the following baseline predictors: age, sex, Pulmonary category at baseline, days since symptom onset, NEW score, and indicator variables for the following risk factors: asthma/COPD, diabetes, CVD, heart failure, hypertension, HIV or other immune deficiency, and renal impairment. The risk score will be derived from the pooled data for the investigational agent/placebo groups. Thus, the risk score will be specific to each investigational agent.

14 Unblinding of Treatment Comparisons

For any investigational agent, trial results will be unblinded when the pre-specified number of primary endpoints is reached; results may be unblinded earlier upon the recommendation of the DSMB if the sponsor and study leadership concur. In this case, trial results for the investigational agent will be unblinded and reported with available data through 90 days of follow-up. After that, data collection will continue as outlined in the data collection plan; under protocol version 3.0, death and re-hospitalizations will be recorded through 18 months.

While the trial is ongoing, access to any data summaries by treatment group (investigational agent or control groups) will be restricted to the members of the DSMB, the DSMB's Executive Secretary, and the unblinded statisticians.

When the trial for an investigational agent is concluded, data for the investigational agent and the corresponding pooled control group will be unblinded and provided to the study team.

The timing of the unblinding of data for one agent may require consideration, if:

- the control group is substantially shared with another agent for which the trial is still ongoing, **and**
- pooled data on treatment outcomes for the ongoing trial are available to investigators.

In this case, the need for a speedy unblinding has to be balanced with maintaining trial integrity for other agents in the platform trial, and the DSMB will be consulted as to the timing of the unblinding.

15 Distribution of Reports

- Open report: ACTIV-3 leadership team; DAIDS Medical Officer; selected NIAID staff; representatives of the companies; and all recipients of the unblinded closed report. After the DSMB meeting, the open report and the DSMB summary statement will be posted to the trial's web site, open to all investigators.
- Closed report: DSMB members, Executive Secretary of the DSMB, unblinded statisticians.
- Web reports (accessible by all investigators and study staff):
    - Enrollment summaries by site and over time (updated daily)
    - Baseline characteristics
    - Selected summary measures on data quality and study conduct (pooled across treatment groups).
- Additionally, selected summary measures on study conduct will be provided to study leadership upon request (pooled across treatment groups).

16 References

1. Whitehead J. Sample size calculations for ordered categorical data. Stat Med 1993;12:2257-71.
2. McCullagh P. Regression Models for Ordinal Data. Journal of the Royal Statistical Society Series B (Methodological) 1980;42:109-42.
3. Aalen OO, Johansen S. An Empirical Transition Matrix for Non-Homogeneous Markov Chains Based on Censored Observations. Scandinavian Journal of Statistics 1978;5:141-50.
4. Gray RJ. A Class of K-Sample Tests for Comparing the Cumulative Incidence of a Competing Risk. Annals of Statistics 1988;16:1141-54.
5. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. J Am Stat Assoc 1999;94:496-509.
6. Zhou B, Latouche A, Rocha V, Fine J. Competing risks regression for stratified data. Biometrics 2011;67:661-70.
7. Division of AIDS (DAIDS) Table for Grading the Severity of Adult and Pediatric Adverse Events, Corrected Version 2.1. 2017. (Accessed August 14, 2020, at https://rsc.niaid.nih.gov/sites/default/files/daidsgradingcorrectedv21.pdf.)
8. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J 2012;33:176-82.
9. Wei LJ, Lin DY, Weissfeld L. Regression-Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. J Am Stat Assoc 1989;84:1065-73.
10. ACTIV-3 TICO LY-CoV555 Study Group, Lundgren JD, Grund B, et al. A Neutralizing Monoclonal Antibody for Hospitalized Patients with Covid-19. N Engl J Med 2021;384:905-14.
11. Lundgren JD, Grund B, Barkauskas CE, et al. Responses to a Neutralizing Monoclonal Antibody for Hospitalized Patients with COVID-19 According to Baseline Antibody and Antigen Levels Annals of Internal Medicine 2021.
12. Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the Treatment of Covid-19 - Preliminary Report. N Engl J Med 2020.
13. Proschan MA, Lan KKG, Wittes JT. Statistical monitoring of clinical trials : a unified approach. New York, NY: Springer; 2006.
14. Logan BR, Zhang MJ. The use of group sequential designs with common competing risks tests. Stat Med 2013;32:899-913.
15. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 1979;6:65-70.

Appendix A. Definition of the Pulmonary and Pulmonary+ ordered categorical outcomes

The Pulmonary categorical outcome is primarily defined based on oxygen requirements. The categories of the Pulmonary+ outcome are similar, except that categories 4 and 5 also capture selected extra-pulmonary complications, highlighted in red below.

Pulmonary outcome	Pulmonary+ outcome
1. Can independently undertake usual activities with minimal or no symptoms	1. Can independently undertake usual activities with minimal or no symptoms
2. Symptomatic and currently unable to independently undertake usual activities but no need of supplemental oxygen (or not above premorbid requirements)	2. Symptomatic and currently unable to independently undertake usual activities but no need of supplemental oxygen (or not above premorbid requirements)
3. Supplemental oxygen (<4 liters/min, or <4 liters/min above premorbid requirements)	3. Supplemental oxygen (<4 liters/min, or <4 liters/min above premorbid requirements)
4. Supplemental oxygen (≥4 liters/min, or ≥4 liters/min above premorbid requirements, but not high-flow oxygen)	4. Supplemental oxygen (≥4 liters/min, or ≥4 liters/min above premorbid requirements, but not high-flow oxygen) or any of the following: stroke (NIH Stroke Scale [NIHSS] ≤14), meningitis, encephalitis, myelitis, myocardial infarction, myocarditis, pericarditis, new onset CHF NYHA class III or IV or worsening to class III or IV, arterial or deep venous thromboembolic events
5. Non-invasive ventilation or high-flow oxygen	5. Non-invasive ventilation or high-flow oxygen, or signs and symptoms of an acute stroke (NIHSS >14)
6. Invasive ventilation, extracorporeal membrane oxygenation (ECMO), mechanical circulatory support, or new receipt of renal replacement therapy	6. Invasive ventilation, ECMO, mechanical circulatory support, new receipt of renal replacement therapy, or vasopressor therapy
7. Death	7. Death

The term “usual activities”, in categories 1 and 2 for both outcomes, refers to activities of daily living that the participant was able to undertake prior to the current illness.

Appendix B. Definition of Clinical Organ Failure and Serious Infection

According to the protocol, section 4.2.2., *clinical organ failure* is defined by development of any one or more of the following clinical events (see PIM for criteria for what constitutes each of these conditions):

a. Respiratory dysfunction:
    1. Respiratory failure defined as receipt of high flow nasal oxygen, non-invasive ventilation, invasive mechanical ventilation, or ECMO

b. Cardiac and vascular dysfunction:
    1. Myocardial infarction (MI)
    2. Myocarditis or pericarditis
    3. Congestive heart failure (CHF): new onset NYHA class III or IV, or worsening to class III or IV
    4. Hypotension requiring institution of vasopressor therapy

c. Renal dysfunction:
    1. New requirement for renal replacement therapy

d. Hepatic dysfunction:
    1. Hepatic decompensation

e. Neurological dysfunction
    1. Acute delirium
    2. Cerebrovascular event (stroke, cerebrovascular accident [CVA])
    3. Transient ischemic events (i.e., CVA symptomatology resolving <24 hrs)
    4. Encephalitis, meningitis or myelitis

f. Haematological dysfunction:
    1. Disseminated intravascular coagulation
    2. New arterial or venous thromboembolic events, including pulmonary embolism and deep vein thrombosis
    3. Major bleeding events (>2 units of blood within 24 hours, bleeding at a critical site [intracranial, intraspinal, intraocular, pericardial, intraarticular, intramuscular with compartment syndrome, or retroperitoneal], or fatal bleeding).

Serious infection is defined as:

g. Serious infection:
    1. Intercurrent, at least probable, documented serious disease caused by an infection *other than* SARS-CoV-2, requiring antimicrobial administration and care within an acute-care hospital.

Appendix C. Safety Data Collection

Table C-1. Overview of Safety Data Collection (protocol version 3, section 10).*

	Infusion +2 hrs	Days 0-7	Day 14	Day 28	Day 90	Months 6, 12, and 18
Infusion-related reactions and symptoms	X					
Clinical AEs of any grade severity		X	X	X		
Grade 3 and 4 clinical AEs from Day 7 through Day 28[1]			X	X		
Targeted laboratory abnormalities of any grade		X (Day 5)				
Hospital admissions and deaths		Collected through Month 18				
Targeted clinical events collected as study endpoints[2]	Collected through Day 90					
SAEs not exempt from reporting (i.e., not considered a protocol specified exempt event)[2]	Collected through Day 90					
Any SAE related to study intervention	Collected through Day 90					
Unanticipated problems	Collected through Day 90					

[1] Grade 3 and 4 clinical AEs on Days 1-7 are reported each day; those occurring between Days 8 and 14 are reported at the Day 14 visit, and those occurring between Days 15 and 28 are reported at the Day 28 visit.

[2] Protocol-specified exempt serious events (PSEE); these events are listed below. See section 10.2.5 of the TICO study protocol and the PIM for information on PSEE.

* In protocol version 2.0, the data collection is identical, except that hospital re-admissions and deaths are collected through Month 12 only.

Protocol-specified exempt events (protocol section 10.2.5)
The following events are protocol-specified exempt events. They are **not** reported as AEs or SAEs, **unless** the investigator considered that there was a reasonable possibility that the study intervention (blinded investigational agent/ placebo or study-supplied SOC treatment) caused the event.

- Death
- Stroke
- Meningitis
- Encephalitis
- Myelitis
- Myocardial infarction
- Myocarditis
- Pericarditis
- New onset of worsening of CHF (NYHA class 3 or 4)
- Arterial or deep vein thromboembolic events
- Respiratory failure defined as receipt of high flow nasal oxygen, non-invasive ventilation, invasive mechanical ventilation or ECMO
- Hypotension requiring vasopressor therapy
- Renal dysfunction requiring renal replacement therapy
- Hepatic decompensation
- Neurologic dysfunction, including acute delirium and transient ischemic events
- Disseminated intravascular coagulation
- Major bleeding events
- Serious infections

Appendix D. Schedule of Assessments

Table D-1. Schedule of Assessments (protocol version 5.1, Appendix B).*

	Screen or Day 0	Day 0	Follow-up Study Day; shaded columns denote in-person visits															
Day	**−1/0**[1]	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**14**	**28**	**42**	**60**	**75**	**90**	**6M**	**12M**	**18M**
Acceptable deviation from day	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**+1**	**+2**	**+3**	**+3**	**+5**	**+5**	**+10**	**±14**	**±14**	**±14**
ELIGIBILITY & BASELINE DATA																		
Informed consent	X																	
Baseline medical (incl. duration of COVID-19) and social history	X																	
Baseline medications	X																	
Symptom-directed physical exam by the clinical team	X																	
Review SARS-CoV-2 test results	X																	
Local laboratory testing	X						X											
Urine pregnancy test or other documentation of pregnancy status	X																	
STUDY INTERVENTION																		
Randomization		X																
Study Drug/Placebo Administration		X																
Assess infusion completion and adverse reactions		X																
STUDY PROCEDURES																		
Clinical assessment for pulmonary ordinal outcome	X	X	X	X	X	X	X	X	X	X	X							
Clinical assessment for pulmonary+ ordinal outcome	X	X	X	X	X	X	X	X	X									
Vital signs for NEW score assessment	X																	
Respiratory function scale assessment	X																	
Hospitalization status					X		X		X	X	X		X		X	X	X	X
Changes in residence/facility										X	X	X	X	X	X			
Interim medical history									X	X	X		X		X			
Interim medications							X				X							
Clinical AEs of any grade on days indicated		X	X	X	X	X	X	X	X	X	X							
Clinical AEs reaching grade 3 or 4 severity through Day 28										X	X							
Research sample storage (plasma and serum)[2]		X	X		X		X				X				X			
Midturbinate swab for central SARS-CoV-2 viral load testing[2]		X																
SAEs and unanticipated problems			Report as they occur															
Deaths			Report as they occur															
Hospitalization Summary			Report upon hospital discharge															
Hospital Readmissions			Report upon hospital discharge															

[1] Screening must be performed within 24 hours of randomization.

[2] Blood draw and swab collection in some cases can be obtained after randomization but before the infusion. If it is not possible to do an in-person on Day 3 or Day 5, the blood draws may be done one day earlier or one day later (but the participant should be telephoned to record the clinical data on the indicated study day).

* In addition, in protocol versions 3.0 and higher, pregnancy outcomes will be collected for female participants who become pregnant during the study.

Appendix E. Modifications to the Statistical Analysis Plan for AZD7442

Investigational Agent

Appendix H4 (version 2 dated 09 April 2021) of the TICO protocol describes agent-specific protocol elements for AZD7442 by AstraZeneca. AZD7442 is comprised of two immunoglobulin G (IgG)-1 kappa monoclonal antibodies (AZD8895 and AZD1061) which neutralize SARS-CoV-2 by binding to unique, non-overlapping epitopes on the RBD of the viral spike protein.

E-1 Analysis population

Treatment comparisons for **efficacy and safety** outcomes will be by modified intention-to-treat (mITT). The mITT analysis is restricted to participants who received a complete or partial infusion of the investigational agent/placebo; participants who did not receive any of the investigational agent/placebo are excluded. The TICO platform protocol specifies ITT for the efficacy analyses "unless otherwise stated", and mITT for safety analyses.

Justification for changing the study population to mITT for efficacy analyses: In the TICO trial of AZD7442, 1455 patients were randomized (ITT population). In preliminary, near final data, of the 1455 randomized participants, 1417 were infused the blinded AZD7442 or placebo. Of the remaining 38 participants, 35 had withdrawn consent prior to infusion (no follow-up data after Day 0 were collected for 34 of the 35), and information was pending on 3 participants as of December 28, 2021. Because the treatment assignment was blinded, the reasons for not receiving an infusion are independent from the treatment assignment. Therefore, the risk of bias due to restricting analyses to the mITT cohort is low. Moreover, follow-up data are not available for most of the participants who did not receive an infusion. With this approach, the risk and benefits of AZD7442 will be evaluated in the same population.

> **Comment:** Investigators were unblinded to summaries of the infusion status of participants, pooled across treatment groups. These analyses were included in the open DSMB reports.

E-2 Increase in Target Number of Primary Events and Sample Size Beyond That Specified in the Master Protocol

The event target and sample size for AZD7442 was increased on August 19, 2021 with a letter of amendment. This increase was based on the results of other trials (the RECOVERY trial and the ACTIV-3/TICO trials of bamlanivimab, VIR-7831 and BRII-196/198) and utilized pooled outcome data for AZD7442 to estimate the revised sample size. The new event target for the AZD7442 trial was 1,228. It was estimated that extending enrolment to September 30, 2021 would achieve a sample size of 1,500 participants, which was estimated to be required to achieve the event target of 1,228. With this event target a RRR of 1.20 could be detected with 90% power at the 0.05 (2-sided) level of significance. It also provided 90% power to detect a RRR of 1.28 for the subgroup of participants who are seronegative for neutralizing SARS-CoV-2 antibodies (nAb) at study entry (see Letter of Amendment of 19 August 2021).

E-3 Power Estimation Prior to Unblinding

On December 28, 2021, prior to unblinding any treatment comparisons for the trial of AZD7442, power was re-assessed based on updated pooled (AZD7442 and placebo groups combined) primary endpoint results. Pooled primary outcome data on December 28, 2021 are given in the table below.

Table E-1. Recovery status of participants in the AZD7442 trial as of December 28, 2021

	1455 Participants Randomized				
	1417 Infused and 38 not infused (35 withdrew consent)				
	nAb negative	**nAb positive**	**Specimen not yet analyzed**	**Specimen not available**	**TOTAL**
Recovery status through Day 90	**N=616**	**N=657**	**N=107**	**N=37**	**N=1417**
Withdrew prior to recovery	2	4	0	0	6
Day 90 window closed and recovery status unknown	17	5	0	1	23
Day 90 window not yet closed* and recovery status unknown	0	1	3	0	4
Achieved sustained recovery	508	586	76	25	1195
Home < 14 days	7	5	0	0	12
Discharged from hospital, but not home	9	3	0	2	14
Hospitalized	12	10	2	1	25
Died before recovery	61	43	26	8	138

* The Day 90 window extends to Day 100.

The mITT study population included 1417 participants. At this time, most of the 90-day follow-up was completed. The last patient was randomized on September 30, 2021 and the visit window for the Day 90 visit extends to 100 days (January 8, 2022).

Overall, the number of primary events was 1,195, 33 events less than the target of 1,228. As the table below indicates, this does not have a material effect on power.

Table E-2. Power estimates in the full study population: Recovery rate ratios (RRR) that can be detected with 80-90% power and 0.05 (2-sided) significance level, assuming 1195 sustained recovery events pooled across treatment groups

Power	**RRR**
0.90	1.206
0.85	1.189
0.80	1.176

Table E-3 below summarizes power estimates for the subgroup of participants who were nAb-negative at study entry. This subgroup is currently 48% of those for whom the baseline nAb status was determined. We anticipate results for 107 additional patients in January 2022. The effect sizes (recovery rate ratio [RRR]) that can be detected with 90%, 85%, and 80% power for different estimates of the final number of primary events for the nAb-negative subgroup are summarized below; as of December 28, 2021, 508 participants in this subgroup had achieved sustained recovery, and 76 participants among those for whom the nAb status is not yet known.

Table E-3. Power estimates in the nAb negative subgroup: RRRs that can be detected with 80-90% power and 0.05 (2-sided) significance level, assuming 542-562 sustained recovery events

Power	RRR (542 events)	RRR (552 events)	RRR (562 events)
0.90	1.321	1.318	1.315
0.85	1.294	1.291	1.288
0.80	1.272	1.269	1.267

In preliminary data in August 2021, 56% of participants were nAb-negative. With the more complete baseline data on December 28, 2021, this percentage is estimated as 48%.

The power estimates given in Tables E-2 and E-3, the decline in the estimated percentage of participants who were nAb-negative, and further consideration of the relative importance of the treatment comparisons in the full study cohort as well as the in the nAb-negative subgroup led to the plan to test both hypotheses (benefit of AZD7442 in the full study population, and benefit of AZD7442 in the subgroup of nAb-negative participants) simultaneously while controlling the family-wise type 1 error rate. This is described in section E-4 below.

E-4 Primary Analysis: Simultaneous Tests in the Overall Cohort and in the nAb-Negative Subgroup, Method and Power Considerations

Primary analysis: Two hypotheses will be tested using mITT analyses. The family-wise type 1 error rate will be controlled at the 0.05 (2-sided) level of significance. The two primary comparisons are:

- AZD7442 versus placebo for time to sustained recovery in the overall mITT cohort; and
- AZD7442 versus placebo for time to sustained recovery in the subgroup of participants in the mITT cohort who were nAb-negative at study entry.

The Holm method will be used to control the family-wise error rate.[15] With this approach, the two primary comparisons are considered on an equal footing (i.e., co-primary hypotheses). With the Holm method, p-values for the 2 hypotheses will be ordered from lowest to highest. The lower of the two p-values will be compared with a 0.025 (0.05/2) 2-sided level of significance. If that test result is significant (i.e., $p \leq 0.025$), the second p-value will be compared with a 0.05 (2-sided) level of significance. This method controls the family-wise type 1 error rate for both comparisons at 0.05 (2-sided).

Power was estimated considering the estimated number of participants with a primary endpoint of sustained recovery for the nAb negative and positive subgroups. Based on the results of recently reported trials mentioned in section E-2, we expect the recovery rate ratio (RRR) for AZD7442 versus placebo to be greater (more favorable for AZD7442 than placebo) for the nAb-negative subgroup than for the nAb-positive subgroup.

The following assumptions were made for the power estimates given in the figure and table below; projections for event numbers are based on Table E-1:

- Sustained recovery by Day 90 will be experienced by an estimated 545 participants in the nAb-negative subgroup. This projection includes the 508 participants who have had experienced the endpoint by December 28, 2021, and 37 additional participants who

experienced a sustained recovery endpoint (either nAb negative participants with unknown recovery status or participants for whom nAb status at baseline had not yet been determined on December 28, 2021).

- In the overall cohort, 1202 participants will achieve sustained recovery; this projection includes the 1195 participants who had achieved sustained recovery as of December 28, 2021, plus a small proportion of participants with as yet unknown outcome status.
- 48% of the study participants would be nAb-negative at study entry

With 545 primary events in the nAb-negative subgroup, power is 0.90 to detect a RRR=1.32 at the 0.05 (2-sided) level of significance. With the above assumptions, and assuming RRR=1.32 for the nAb-negative subgroup, power for the comparison of AZD7442 and placebo in the overall cohort and within the nAb-negative subgroup was estimated for a range of values of the RRR for the nAb-positive subgroup. In Table E-4 below, the first column shows the assumed RRR among the nAb-positive subgroup, the second column shows the corresponding overall RRR (assuming RRR=1.32 among the nAb-negative subgroup), and columns 3 and 4 show the estimated power to detect RRR=1.32 among the nAb-negative subgroup and the computed RRR (column 2) in the overall study cohort, respectively, using Holm's method to control the simultaneous Type 1 error to 0.05. For example, assuming RRR=1.0 among the nAb-positive subgroup and RRR=1.32 among the nAb-negative subgroup, the overall RRR would be 1.14, and the treatment difference between AZD7442 and placebo would be detected with power of 0.858 in the nAb-negative subgroup, and with power of 0.63 in the overall study population.

Comment: In a single test with 5% significance level, power was estimated at 90% to detect RRR=1.32 in the nAb-negative subgroup; assuming RRR=1.14 in the full study cohort, the power in the nAb-negative subgroup is decreased to 86% with Holm's procedure because the p-value for the nAb-negative subgroup would more likely be the smaller of the two p-values, and thus would be compared to 0.025 instead of 0.05.

Table E-4. Power to detect a treatment difference in sustained recovery between AZD7442 and placebo within the nAb-negative subgroup and overall, assuming an RRR of 1.32 among the nAb-negative subgroup and a range of RRRs between 1.30 and 0.95 for the nAb-positive subgroup, and using Holm's method to control the familywise Type 1 error.

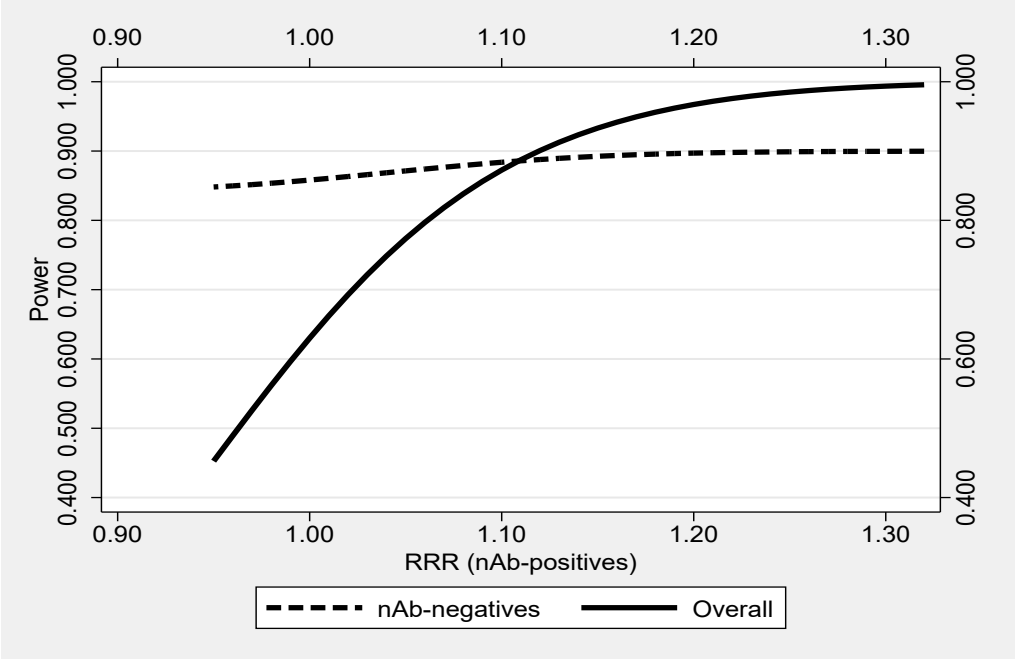
Assumed RRR		**Power using Holm's method**	
Among the nAb-positive subgroup	Overall*	To detect RRR=1.32 among the nAb-negative subgroup	To detect the overall RRR (column 2)
1.30	1.31	0.900	0.994
1.25	1.28	0.899	0.985
1.20	1.26	0.897	0.967
1.15	1.23	0.892	0.933
1.10	1.20	0.884	0.873
1.05	1.17	0.871	0.774
1.00	1.14	0.858	0.630
0.95	1.11	0.848	0.453

* RRR for sustained recovery in the overall study population, assuming RRR=1.32 among the nAb-negative subgroup and the RRR in column 1 among the nAb-positive subgroup

Figure E-1 below shows the data from Table E-4 in a graph. The power of Holm's testing procedure to detect a treatment effect is plotted versus a range of assumed RRRs among the nAb-positive participants (horizontal axis, column 1 of Table E-4); the dashed line shows the

power to detect RRR=1.32 for the nAb-negative subgroup (column 3 in Table E-4), and the solid line shows the power to detect the corresponding treatment effect in the overall cohort (column 4 in Table E-4).

Figure E-1. Estimated power to detect a treatment difference in sustained recovery between AZD7442 and placebo in the overall study population (solid line) and within the subgroup of nAb-negative participants (dashed line), for a range of RRRs among nAb-positive participants and assuming RRR=1.32 for nAb-negative participants.



E-5 Data Analysis Considerations for the Two Co-Primary Hypotheses

The analysis of the primary endpoint will follow the methods described in sections 3 (stratification) and 8.1 (Gray's test for the primary comparisons and Fine-Gray models to estimate RRRs and 95% CIs, stratified by study site pharmacy). The following additional considerations apply to the AZD7442 study:

- The same stratification will be used for testing both co-primary hypotheses. Using blinded baseline data available on January 21, 2022, the strata were determined as follows. To avoid sparse strata, study pharmacies were pooled following the principles described in section 3. Seven strata, each including at least 20 participants in the nAb-negative subgroup, were formed using large sites where possible and combining small sites with each other or with a larger site where that was not. In addition, some very small sites that only enrolled participants who were nAb-positive were combined with one of the 7 strata. For the 99 sites that enrolled at least one participant, the 7 strata will be defined as follows: 1) 3 strata for the 3 U.S. sites that enrolled at least 20 nAb negative participants; 2) 1 stratum for all other U.S. sites; 3) 1 stratum for sites in Denmark, Greece, the United Kingdom, and Singapore; 4) 1 stratum for sites in Spain; and 5) 1 stratum for sites in Uganda.

- No participants were randomized in disease severity stratum 2 (defined in section 1.2; in particular, stratum 2 included participants requiring invasive mechanical ventilation or ECMO). Therefore, the originally planned stratification of analyses by disease severity stratum can't be carried out. Early in the trial, participants requiring non-invasive ventilation (NIV) or oxygen from a high-flow nasal cannula (HFNC) also were excluded. Beginning July 19, 2021, eligibility was expanded to include participants requiring NIV or HFNC. Considering this expansion of eligibility criteria, a sensitivity analysis will be carried out using a Fine-Gray model that only stratifies by oxygen requirements at baseline. A sensitivity analysis without any stratification will also be carried out.
- A subgroup analysis by calendar date of enrollment will also be carried out (2 subgroups: participants enrolled before July 19, 2021 and at or after July 19, 2021). The cut-point was chosen because on July 19, 2021, the first patient who required HFNC or NIV was enrolled after eligibility was expanded to include such patients.

If the primary comparison of AZD7442 versus placebo for time to sustained recovery shows a significant treatment effect in nAb-negative participants, then major safety, efficacy, and subgroup analyses will be conducted for the nAb-negative participants as well as for the full study cohort.

E-6 Interim Monitoring Guidelines

There are 2 changes to the interim monitoring guidelines described in Section 10.3 of this SAP:

- The information fraction for the Lan-DeMets spending function will be based on the target of 1,228 sustained recoveries, not 843.
- As part of the broad assessment of risk/benefit, the DSMB should consider the findings in nAb negative and nAb positive subgroups. For participants who are nAb negative, greater benefit with the investigational agent is hypothesized.

  Comment: The antibody status at baseline was unknown for a large proportion of participants until late in the trial. Therefore, separate interim monitoring boundaries were not planned for this subgroup.

E-7 Subgroup Analyses by Vaccine and Immunosuppression Status

The hypotheses in this subsection were developed for the comparison of AZD7442 (a long-acting antibody combination) versus placebo, but apply more generally to monoclonal antibodies and antiviral agents. The outcomes, and the specific statistical models used for the subgroup analyses for each of the outcomes were described in Section 9.1.

The subgroup analyses by vaccination status and immunosuppression status address the following hypothesis:

Hypothesis 1: AZD7442 will have greater efficacy in non-vaccinated, compared to fully or partially vaccinated, study participants.

Hypothesis 2: AZD7442 will have greater efficacy in immunosuppressed, compared to not immunosuppressed, study participants.

For vaccination status, three mutually exclusive subgroups will be formed according to the current CDC definition of vaccination status:

a. **Fully-vaccinated:** Subjects who have onset of acute symptoms of SARS-CoV-2 infection at least 14 days after receiving:
   i. the **second** dose of a two-dose SARS-CoV-2 vaccine series, OR
   ii. a **single** dose of the Janssen (J&J) SARS-CoV-2 vaccine

b. **Partially-vaccinated:** Subjects who have:
   i. received **only one dose** of two-dose SARS-CoV-2 vaccine series, OR
   ii. onset of acute symptoms of SARS-CoV-2 infection prior to 14 day following the administration of the **second** dose of a two-dose SARS-CoV-2 vaccine series or a **single** dose of the Janssen (J&J) SARS-CoV-2 vaccine

c. **Non-vaccinated:** Never received any dose of a SARS-CoV-2 vaccine.

Comment: Information about booster doses is not available.

Subgroup analyses by immunosuppressive status will consider three mutually exclusive subgroups, as defined below:

- **Immunosuppressed**, defined as any of the following**:**
  - Receiving antirejection medication (after solid or stem cell transplant)
  - Receiving biologic medicine to treat autoimmune disease or cancer, excluding IL-1 inhibitors, IL-6 inhibitors, Janus kinase inhibitors, and TNF inhibitors
  - HIV or other immunosuppressive condition
- Receiving any of the following immunomodulators: IL-1 inhibitors, IL-6 inhibitors, Janus kinase inhibitors, or TNF inhibitors, but not “immunosuppressed” as defined above
- None of the above

Data from TICO and other studies have shown that baseline vaccination status and immunosuppressive status are correlated; in particular, in our study population of patients who are hospitalized with COVID-19 pneumonia, the proportion of immunosuppressed patients is higher among those who are fully *vaccinated* compared with those who are unvaccinated. In addition, studies have found that the decline over time in immunity varies between vaccines. In order to address potential confounding between vaccination status, immunosuppressive status, and type of vaccine, models for subgroup analyses by vaccination status will include immunosuppressive status and type of vaccine (mRNA versus other) as covariates. Similarly, models for subgroup analyses by immunosuppressive status will include vaccination status and type of vaccine as covariates.

The following **exploratory analyses** will be conducted if the sample size is large enough for meaningful analyses:

- Subgroups will also be formed by combinations of *vaccination status and vaccine type* (4 categories: not vaccinated; partially vaccinated; fully vaccinated with non-mRNA vaccine; fully vaccinated including at least one mRNA vaccine in the sequence). Models for this subgroup analysis will be adjusted for immunosuppressive status.

- Subgroups will be formed by the bivariate combinations of *vaccination status and immunosuppressive status*.

In order to facilitate the interpretation of the results of the subgroup analyses, baseline levels of antibodies, antigen, and SARS-CoV-2 RNA will be summarized by vaccination status and by immunosuppressive status. These biomarkers include the following:

- SARS-CoV-2 antibody levels (nAb by GenScript, anti-N by BioRad, and anti-Spike IgG antibody levels by Quanterix assays)
- SARS-CoV-2 antigen levels in plasma
- Viral RNA by midturbinate swabs

See Appendix F for details on assays and outcome measures.

E-8 Subgroup Analyses by Viral Strain

Between March 1 and September 30, 2021, 1455 participants were randomized to AZD7442 or placebo. The enrollment period included time periods both prior to and after the emergence of the Delta variant. Higher levels of plasma nucleocapsid antigen at baseline were observed during enrollments in July-September compared to enrollments in March-June, which coincided with the emergence and global replacement of the Alpha/other viral variants with the Delta variant. Observational studies suggest worse clinical outcome from being infected with the Delta than the Alpha viral variant.

Research Question 1: Does the treatment effect of AZD7442 versus placebo differ based on the viral strain that the participant is infected with?

Hypothesis: The benefit of AZD7442 differs between participants who were infected with the Delta variant versus those who were infected with other viral variants.

Justification: The AZD7442 nMAb combination binds to non-overlapping epitopes on the SARS-CoV-2 Spike protein, presumably decreasing viral replication and dissemination by blocking the ability of the Spike protein to bind to host cells expressing ACE2. As such, participants who were infected with the Delta variant versus those who were infected with other viral variants may benefit more from treatment with ADZ7442. This may be particularly pronounced in the subset of participants who have not yet mounted their own endogenous neutralizing antibody response.

An alternate hypothesis is that the dose of AZD7442 is better able to contain viral replication among persons infected with a less virulent viral strain, and hence the relative benefit from using AZD7442 could be more pronounced among participants infected with non-Delta than Delta viral variants.

Statistical analyses:
To address the hypothesis of a differential treatment effect, the investigational agent will be compared versus placebo in subgroups by SARS-CoV-2 variant (Delta versus other), and tests for differential treatment effects across the subgroups will be performed by testing for

interactions between the treatment and subgroup indicators. Subgroup analyses will be performed for time to sustained recovery (primary efficacy outcome) and key safety outcomes, including the composite outcomes of *grade 3 or 4 AEs, SAEs, organ failure, serious infections, or death* through Day 28; *SAEs, organ failure, serious infections, or death* through Day 90; and death through Day 90. Statistical methods for the subgroup analyses are described in Section 9.1.

Because the composition of the study population may have changed over time, including an increase in the proportion of participants infected with the Delta variant, an increase in the proportion of vaccinated persons in the general population, and changes in the population with respect to supplementary oxygen requirements, models for these subgroup analyses will also consider adjustment for vaccination status, and for supplementary oxygen requirements at study entry (categories of the ordinal pulmonary outcome).

Research Question 2: What is the effect of specific mutations?

AZD7442 is a combination therapy targeting non-overlapping epitopes in the RBD. While mutations found in variants can impact the binding of either one of the nMAbs (including at positions R346, K444, V445, G446, N450, L452 G476, E484, Q493K) the combination is hypothesized to retain activity. Nevertheless, high quality sequencing of the Spike gene will provide information on any relevant RBD mutations that could impact binding of one or both nMAbs.

Hypothesis: The relative benefit of AZD7442 (vs placebo) is larger for those infected with virus without mutations at binding-relevant positions compared with those with such virus mutations.

Statistical Analyses:
It is expected that the rate of viral variants than contains mutations at the binding-relevant positions is low, with correspondingly low power for statistical tests. For individual mutations that occur at sufficiently high rates, and pooled across mutations, subgroup analyses will be performed comparing AZD7442 versus placebo within subgroups by the presence versus absence of the mutations. The treatment effect will be compared across subgroups by testing for an interaction between the treatment and subgroup indicators. Subgroup analyses will be performed for time to sustained recovery and key safety outcomes, as described above; statistical methods are described in section 8.4 of this SAP. Models will include vaccination status and supplementary oxygen requirements (categories of the ordinal pulmonary outcome) at baseline as covariates.

These are exploratory analyses. In addition, if mutations on other sites occur with sufficient frequency, similar analyses will be performed to identify mutations that are associated with the outcome of the treatment.

E-9 Missing Outcomes and Missing Baseline nAb Status

Missing primary outcome results and 90-day survival status due to withdrawal of consent or loss to follow-up will be summarized by treatment group, and tests will be performed whether the proportion of missing data is independent of the treatment assignment. Baseline factors related to these missing outcome data, overall and for each treatment group, will be determined using logistic regression. Candidate baseline factors for this investigation include, but are not limited to, categories of oxygen requirement (pulmonary outcome), antigen levels, and the disease progression risk score (section 13.2). Summaries of missing outcome data will also include when patients with missing data were last seen and their last known clinical status (e.g., discharged, hospitalized, alive at home).

As sensitivity analyses, multiple imputation may be used to estimate missing outcomes. Other sensitivity analyses may also be considered following unblinding when the final extent of missing data is known and the pattern of missing data is investigated. As of December 28, 2021, 6 participants had withdrawn consent and 23 additional patients had unknown sustained recovery status at 90 days. Follow-up is ongoing, including attempts to reach patients who are lost to follow-up. Thus, these numbers are likely to change.

We anticipate that 30-60 participants will not have a baseline specimen that can be used to determine nAb status. The primary analysis in the nAb-negative subgroup will be restricted to participants with known (nAb-negative) status, while the total cohort will include both, participants with known and unknown antibody status. As a sensitivity analysis, we will impute nAb status for these participants using baseline characteristics that include age, oxygen requirements, duration of time between symptom onset and randomization, and SARS-CoV-2 vaccination status. Using a logistic regression model, baseline predictors of nAb status will be determined. The coefficients from the regression model will be used to classify the participants with missing nAb status as negative or positive. For the imputation of the missing nAb status, a range of cut-points will be used: participants will be imputed as nAb-negative if the predicted probability of having been nAb-negative at baseline is 50% or greater, 60% or greater, and 70% or greater. The primary comparison for the sustained recovery endpoint in the nAb-negative subgroup will be repeated including both those with known nAb-negative status as well as those who were imputed as nAb-negative, in separate analyses for each of the three imputation cut-points.

Appendix F. Antibody, Antigen, and viral RNA Assays and Outcomes

SARS-CoV-2 antibody and antigen levels were measured in plasma specimens collected at baseline (Day 0), Day 1, Day 3, and Day 5. Antibody and antigen levels were determined centrally, by the Frederick National Laboratory, blinded to treatment group.

Antibody Levels
SARS-CoV-2 plasma antibody levels were measured using two assays:

- Levels of neutralizing antibodies (**nAbs**) directed against the SARS-CoV-2 receptor binding domain (RBD) were determined using the **GenScript** SARS-CoV-2 Surrogate Virus Neutralization Test (sVNT) assay (GenScript, Piscataway, NJ, USA).
    - Levels of nAbs reported as "percent binding inhibition". Specimens with levels <30% are considered nAb negative, levels $\geq$ 30% are considered positive for nAbs (30% is the manufacturer's cutoff for positivity).

- **BioRad** Platelia SARS-CoV-2 Total Ab assay (BioRad, Hercules, CA, USA), measuring total (IgA, IgG, and IgM) anti-nucleoprotein (**Anti-N**) antibodies.
    - Results of this antibody measurement are reported as "specimen ratios". Specimen ratios are defined as the specimen optical density (OD) divided by the optical density of the cut-off control R4 (OD_MR4). According to the manufacturer, specimen ratios less than 0.8 are considered negative, those with a specimen ratio between 0.8 and 1.0 are considered equivocal, and those $\geq$ 1.0 are considered positive for the presence of antibodies.

    - Equivocal BioRad antibody levels will be combined with negative levels for all analyses unless otherwise stated.

- **Quanterix** anti-Spike IgG semi-quantitative antibody assay (Simoa® Semi-Quantitative SARS-CoV-2 IgG Antibody Test, Quanterix, Bellerica, MA, USA).
    - Automated paramagnetic microbead-based immunoassay intended for qualitative and semi-quantitative detection of IgG antibodies to SARS-CoV-2 in human dipotassium EDTA plasma using the Quanterix HD-X immunoassay system.
    - Functional LLoQ of 82 ng/mL
    - Antibody-positive defined as $\geq$ 770 ng/mL per manufacturer's recommendation ("clinical cut-off")
    - The performance of the assay has not been established in individuals that have received a COVID-19 vaccine

Comment: For the purpose of comparing natural immunity and whether nMAbs have potential benefit in those with such, the GenScript assay will be used. This attempts to quantify neutralizing antibody titers, whereas the BioRad assay captures any type of anti-N antibody (against a section of virus not causing neutralization).

Plasma Antigen Levels
SARS-CoV-2 nucleocapsid antigen levels were determined in 90 µL plasma in duplicate using

a **Quanterix** assay (Simoa® SARS-CoV-2 N Protein Advantage, Quanterix, Bellerica, MA, USA). The lower limit of quantification for the assay is 3 ng/L.

- Antigen levels < 3 ng/L are considered "antigen negative".

Comments:

- When analyzed as continuous variable, Quanterix antigen levels <3 will be set to 2.9 ng/L.
- When analyzed as continuous variable, antigen levels will be log_{10}-transformed.

SARS-CoV-2 Viral RNA was assessed at baseline from mid-turbinate nasal swabs. Qualitative and quantitative assessments of the SARS-CoV-2 RNA in viral transport media (proxy for viral load) by RT-PCR were made centrally by ABML.

- *Qualitative RT-PCR analysis:* Extraction, master mix preparation, and RT-PCR were performed as described in the CDC 2019-Novel Coronavirus Real-Time RT-PCR Diagnostic Panel. RT-PCR was performed on an Applied Biosystems QuantStudio 7 Flex. Ct scores <40 for both nCoV N1 and nCoV N2 probe sets are scored as positive for the presence of SARS-CoV-2 RNA.
- *Quantitative RT-PCR analysis:* Quantitative RT-PCR analysis of the samples used the same RNA extracts prepared for the qualitative assay. Assay conditions were the same as outlined in the CDC protocol except the RNaseP probe was not used.

Comments:

- For subgroup analyses, RNA levels in viral transport media will be categorized as (<10,000 cp/mL [low] versus >10,000 cp/mL [high]); the cut-point is close to the median viral RNA level in the TICO bamlanivimab trial. Indeterminate levels will usually be combined with negative levels for analyses.
- When analyzed as continuous variable, RNA levels will be log_{10}-transformed.

Appendix G. Modifications to the Statistical Analysis Plan for MP0420

Investigational Agent
Appendix H5, version 2.0 (09 April 2021) of the TICO protocol describes agent-specific protocol elements for MP0420 which is being developed by Molecular Partners and Novartis. MP0420 is multi-valent DARPin® molecule consisting of 5 DARPin® domains. MP0420 simultaneously and specifically binds to 3 epitopes of the receptor-binding domain of the SARS-CoV-2 spike protein with three different domains and to human serum albumin with two domains.

Analysis Population
Treatment comparisons for efficacy and safety outcomes will be by modified intention-to-treat (mITT). The mITT analysis is restricted to participants who received a complete or partial infusion of MP0420/placebo; participants who did not receive any MP0420/placebo are excluded. The TICO platform protocol specifies ITT for the efficacy analyses "unless otherwise stated", and mITT for safety analyses.

Justification for changing the study population to mITT for efficacy analyses: In the TICO trial of MP0420, 496 participants were randomized (ITT population). Of these, 485 were infused (complete or partial) with MP0420/placebo and 11 were not. Of the 11 participants not infused, 9 withdrew consent prior to infusion. Follow-up data after Day 0 was collected for the 2 participants not infused, but remained consented. Because the treatment assignment was blinded, the risk of bias due to restricting analyses to the mITT cohort is low. With this approach, the risk and benefits of MP0420 will be evaluated in the same population.

> **Comment:** Investigators were unblinded to summaries of the infusion status of participants, pooled across treatment groups. These analyses were included in the open DSMB reports.

Hypersensitivity Events of Special Interest
In October 2021, we were informed of 3 participants who experienced hypersensitivity reactions in a Molecular Partners study of 20 healthy volunteers. The 3 participants received a 400 mg subcutaneous dose of MP0420 and each developed non-serious (mild to moderate) rash and joint/body pain 10 to 20 days after the single subcutaneous infusion. A total of 8 participants received the 400 mg dose in this study.

These cases were discussed on October 20, 2021, at a TICO investigator meeting and are also described in a memorandum from trial leadership to investigators in ACTIV-3 (TICO) on November 2, 2021. At that time we requested that investigators increase vigilance of participants who developed rash and simultaneously joint, muscle, and/or body aches.

On November 5, 2021, revised informed consents and drug information sheets that described these events were sent to sites and to the Advarra central IRB.

On December 7, 2021, all participants for whom rash adverse events had been reported through Day 28 were identified. Sites were asked to complete an eCRF for each event. Eleven participants experienced rash events through Day 28. The number and percentage of participants with a rash event by treatment group will be summarized. Rash events associated with myalgia, arthralgia, or generalized aches and pains will also be summarized by treatment group.

Line listing of events will include severity grade, rash characteristics, location, and other solicited events that occurred simultaneously.

No formal statistical testing will be carried out on the rash events alone. In a supplemental analysis for the Day 28 composite safety outcome which includes deaths, SAEs, organ failure, serious infection and grade 3 or 4 adverse events, the rash adverse events will be included as a separate component.

Subgroup Analyses
In addition to the subgroup analyses described in section 9, those described in section E-7 (vaccine and immunosuppressive status) and E-8 (viral strain) of Appendix E will be carried out for MP0420 for the outcomes listed in section 9.

Appendix H. List of Acronyms

ACTIV	Accelerating COVID-19 Therapeutic Interventions and Vaccines
ACTT	Adaptive COVID-19 Treatment Trial
ADE	Antibody-dependent enhancement
AE	Adverse event
Anti-N	Anti-nucleoprotein
ARDS	Acute respiratory distress syndrome
CHF	Congestive heart failure
CI	Confidence interval
CIF	Cumulative incidence curve
CMH	Cochran-Mantel-Haenszel [test]
COVID-19	Coronavirus-Induced Disease 2019
CVA	Cerebrovascular accident
DSMB	Data and Safety Monitoring Board
ECMO	Extracorporeal membrane oxygenation
EU	European Union
FDA	Food and Drug Administration (US)
GCP	Good Clinical Practice
GDPR	General Data Protection Regulation
GEE	Generalized estimating equations
GMT	Geometric mean titer
HR	Hazard ratio
ICC	International Coordinating Center
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICU	Intensive care unit
IgG	Immunoglobulin G
IL-1	Interleukin 1
IL-6	Interleukin 6
INSIGHT	International Network for Strategic Initiatives in Global HIV Trials
IQR	Interquartile range
IRB	Institutional Review Board
ITT	Intention-to-treat
IV	Intravenous
mAb	Monoclonal antibody
MedDRA	Medical Dictionary for Regulatory Activities
MI	Myocardial infarction
mITT	modified intention-to-treat
mL	Milliliter
nAb	Neutralizing antibody, here measured by a GenScript assay
NEW	National Early Warning [score]
NIAID	National Institute of Allergy and Infectious Diseases, NIH (US)
NIH	National Institutes of Health (US)
NIHSS	National Institutes of Health Stroke Scale/Score
NYHA	New York Heart Association
nMAb	Neutralizing Monoclonal Antibodies
OR	Odds ratio
PCR	Polymerase chain reaction

PIM	Protocol Instruction Manual
PT	Preferred term
PSEE	Protocol-specified exempt (serious) events
RNA	Ribonucleic acid
RR	Rate ratio
RRR	Recovery rate ratio
SAE	Serious adverse event
SARS-CoV-1	Severe acute respiratory syndrome coronavirus 1
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SAP	Statistical analysis plan
sHR	Sub-distribution hazard ratio; subhazard ratio
SOC	Standard of care
SUSAR	Suspected unexpected serious adverse reaction
TOC	Trial oversight committee
UMN	University of Minnesota
UP	Unanticipated problem
U.S.	United States of America
WHO	World Health Organization