Supplemental file 3 – Statistical analyses and definitions of endpoints

Efficacy Analyses

Efficacy analyses will be performed and presented separately for each population. Formal analyses (incl. interim analyses), hypothesis testing and descriptive analyses will be performed on data from the Main Trial Population, whereas only descriptive statistics (including an estimate of the treatment effect, two-sided 95% confidence intervals and one-sided p-value) will be presented for the 'moderate to severe' CKD and COVID-19 populations.

Efficacy analyses will be performed for the mITT analysis set, which is defined as all patients in the population who are randomly assigned to a trial drug and for whom administration of trial drug was started. Patients will be analyzed according to the treatment to which they were randomly assigned, regardless of whether they received what was assigned. Strata as entered in the IRT will be used. The mITT analysis set will be considered as the primary analysis set for the efficacy analyses. Enrolment continues during interim analyses hence if the trial is stopped early due to futility or success, the mITT will include overrun patients.

Analysis of the Primary Efficacy Endpoint

The primary efficacy endpoint is "28-day all-cause mortality", defined as the probability to die (from any cause) from the date of randomization up to and including Day 28. The primary analysis will be based on a logistic regression model with site as a random effect, treatment as a fixed effect, mSOFA score as an ordered categorical variable and eGFR at Baseline as the single continuous covariate.

The treatment effect of recAP compared with placebo will be expressed as an adjusted odds ratio together with the 95% confidence interval (CI). The null hypothesis is that the odds ratio is equal to unity and this will be tested against the one-sided alternative hypothesis that the odds ratio is less than unity at an overall one-sided 0.025 significance level. The ordered categorical variable for mSOFA will break the scale at 9 (≤9 versus >9) to comply with the stratified randomization, but further cuts will be considered at the blind review stage, one below 9 and one above 9, based on approximately equal increments in expected mortality on the logit scale for the two treatment groups combined.

It is expected that there will be only a small amount of missing data on Day 28 mortality. Nonetheless, to account for such data, the primary analysis (including the interim looks) will utilize multiple imputation based on a logistic regression model fitted to the group of patients with data on the primary endpoint.

The following sensitivity analyses will be conducted for the primary endpoint:

• Logistic regression as for the primary analysis with the additional covariates APACHE II score and time from fulfilling both inclusion criteria 3 and 4 to time to treatment (hours). This analysis assesses the robustness of the findings to imbalances in those two covariates.

• Day 28 all-cause mortality obtained based on Kaplan-Meier (KM) curves for time to death up to Day 28. The KM curves will be compiled separately for the mSOFA categories (≤9 versus >9) and treatment differences in Day 28 survival rates combined using a stratified z-test. In this analysis, patients with unknown survival status who withdraw prior to Day 28 will be censored at the time of withdrawal. Patients ongoing in the trial who are known to be alive beyond Day 28 at the time of the analysis will be censored at Day 28. Patients LTFU prior to Day 28 will be censored at their last date known to be alive. This analysis assesses the impact of missing data on survival status on Day 28.

• A tipping point analysis in which all recAP patients with missing data on survival status on Day 28 will be considered as being dead, while all placebo patients with missing data on survival status on Day 28 will be considered as being alive, and all possible combinations of missing data between these two extremes will be considered.

The primary endpoint will be analyzed using the **actual** main trial population, defined based on the pre-AKI reference eGFR (≥45 mL/min/1.73 m^2) provided in the **eCRF** in order to assess the impact of mis-allocations. In the primary analysis mentioned above in Section 7.8.1, population assignment was taken from IRT.

Subgroup analyses will be performed for the following subgroups:
• Region (US/non-US)
• Age (<65, ≥65 years)
• Gender
• mSOFA (<=9, >9)
• Other regions (North America, Europe/Australasia, Japan)
• Proven/confirmed infection type at Baseline (proven at Baseline, suspected at Baseline + confirmed later and suspected at Baseline + not confirmed later)
These analyses will use the same model as for the analysis of the primary endpoint in the mITT population, restricted to the respective subgroups. Forest plots of subgroup-specific 2-sided 95% CIs for the difference in 28-day all-cause mortality probabilities will be presented.

Analysis of Secondary Efficacy Endpoints
Multiplicity for the analysis of secondary efficacy endpoints will be controlled by:
• Initiating the test procedures for secondary efficacy endpoints only if the null hypothesis for the primary efficacy endpoint has been rejected, and
• Using sequential conditional testing of null hypotheses for secondary efficacy endpoints in the order as indicated earlier; the nominal 1-sided significance level used within the sequential testing will be at the same alpha allocated to the primary endpoint at the time of the analysis.
If the trial is stopped early the secondary analysis will be repeated in the interim analysis set (excluding sensitivity analyses), using the nominal 1-sided significance level allocated to the primary endpoint at the time of the interim analysis.

Major Adverse Kidney Events (MAKE) 90
MAKE 90 is defined as dead by Day 90 *or* on RRT at Day 90 *or* ≥25% decline in eGFR on both Day 28 and Day 90 relative to the known or assumed pre-AKI reference level. The primary analysis will be based on a logistic regression model with site as a random effect, treatment as a fixed effect, mSOFA score as an ordered categorical variable, and pre-AKI reference eGFR as the single continuous covariate. The presentation of the results and the handling of missing data will be as described for the primary endpoint.
The following sensitivity analysis will be conducted for this endpoint:
Logistic regression as for the primary analysis with the additional covariates APACHE II score and time from fulfilling both inclusion criteria 3 and 4 to time of treatment (hours). This analysis assesses the robustness of the findings to imbalances in those two covariates.
MAKE 90 excluding death due to COVID-19 infection following initial hospital discharge will be analyzed as per the primary analysis logistic regression model which includes site as a random effect, treatment as a fixed effect, mSOFA score as an ordered categorical variable, and pre-AKI reference eGFR as the single continuous covariate.

Days Alive and Free of Organ Support Through Day 28 and Days Alive and Out of the ICU Through Day 28
Days alive and free of organ support through Day 28 is defined as the number of days alive with no MV, RRT, vasopressors or inotropes, and with death within 28 days counting as zero days. Days alive and out of the ICU through Day 28 is defined as the number of days alive and out of the ICU through Day 28 with death within 28 days counting as zero days.
For both of these secondary efficacy endpoints, it is likely that these data will have distributions in each of the two treatment groups that are non-Normal and to deal with this the primary analyses will utilize a non-parametric method. The method for assessing statistical significance will be a re-randomization test comparing the treatment median values for days alive and free of organ support, respecting randomization according to site and mSOFA score. 95% CIs for the difference in the medians will be constructed to aid interpretation.
The following sensitivity analyses will be conducted for these endpoints:
Analysis will be repeated for the subset of patients alive on Day 28 in order to separate out the effect of mortality during the first 28 days.

Time to Death Through Day 90
All-cause mortality up to and including Day 90 will be based on time from date of randomization to date of death. Patients with unknown survival status who withdraw prior to Day 90 will be censored at the time of

withdrawal. Patients known to be alive on Day 90 will be censored on Day 90 and patients LTFU prior to Day 90 will be censored at their last date known to be alive.
Survival curves for each trial drug group will be estimated by the KM method. Median survival and corresponding 2-sided 95% CIs will be computed by the Brookmeyer and Crowley method (Brookmeyer R, Crowley JA. Confidence interval for the median survival time. Biometrics 1982; 38: 29-41).
The primary analysis of this secondary endpoint will be based on the Cox proportional hazards model with site as a random effect, treatment as a fixed effect, mSOFA score as an ordered categorical variable and eGFR at Baseline as the single continuous covariate. The treatment effect will be expressed as a hazard ratio (HR) together with a 95% two-sided CI.
The following sensitivity analyses will be conducted for this endpoint:
• Cox proportional hazards model as for the primary analysis with the additional covariates APACHE II score and time from fulfilling both inclusion criteria 3 and 4 to time of treatment (hours). This analysis assesses the robustness of the findings to imbalances in those two covariates.
• Day 90 all-cause mortality obtained based on KM curves for time to death up to Day 90. The KM curves will be compiled separately for the mSOFA score categories ≤9 and >9 subgroups and treatment differences in the Day 90 survival rates combined using a stratified z-test. This analysis makes no assumptions about the proportionality of the hazard rates and assesses the robustness of the findings to that assumption.
• Mortality at Day 90 excluding death due to COVID-19 infection following initial hospital discharge will be analyzed as per the primary analysis Cox proportional hazards model with site as a random effect, treatment as a fixed effect, mSOFA score as an ordered categorical variable and eGFR at Baseline as the single continuous covariate.

Analysis of Exploratory Efficacy Endpoints
Descriptive statistics will be provided for the following exploratory efficacy endpoints by trial drug group for the mITT sets of the main trial population, 'moderate to severe' CKD population and COVID-19 population and, if applicable, a nominal 1-sided significance level of 0.025 will be used and/or 2-sided 95% CIs will be provided..

• Time to death through Day 180.
• Change in total and individual organ failure scores through Day 7 (based on the mSOFA scores).
• Days alive and free of RRT through Day 28 (with death within 28 days counting as zero days).
• MAKE 28: dead by Day 28 *or* on RRT at Day 28 *or* ≥25% decline in eGFR on both Day 7/ICU discharge (whichever comes first) and Day 28 relative to the known or assumed pre-AKI level.
• Patients alive and free of AKI on Day 7/ICU discharge (whichever comes first) and on Day 28.
• Patients alive and free of new onset CKD or worsening of CKD (defined as any increase in CKD Stage) on Day 90.
• Days alive and free of vasopressor and inotropes through Day 28 (with death within 28 days counting as zero days).
• Days alive and free of MV through Day 28 (with death within 28 days counting as zero days).
• Days alive and out of the hospital through Day 90 (with death within 90 days counting as zero days).
• Incidence of at least one rehospitalization at any hospital through Day 90.
• Change in index values and VAS score based on the EQ-5D-5L questionnaire at Day 28, Day 90 and Day 180.
• The urinary levels of purines through Day 4 at selected sites.

Further exploratory endpoints are defined for each population, as applicable to that population. These exploratory endpoints are pre-defined (i.e., defined prior to the trial being unblinded for the first safety data review) in the trial Statistical Analysis Plan.

Analysis of Safety Endpoints
Safety parameters will be evaluated on the Safety Analysis Set, which is defined as all patients in the population who received any trial drug. Patients will be analyzed according to the first infusion of trial drug they actually received.
Incidence of AEs, SAEs and TEAEs categorized by MedDRA System Organ Class (SOC) and Preferred Term (PT) will be summarized by trial drug group. Adverse event seriousness, severity, relationship to trial drug, and whether leading to discontinuation of trial drug will also be displayed in summaries and listings.
Local laboratory assessments will be summarized using descriptive statistics by trial drug group. Changes from baseline laboratory assessments will be summarized per trial drug group.

Anti-recAP antibodies results will be listed, including the results of the screening test and, if needed, the results of the confirmatory test and titer determination per dose group.

Other Analyses
Information from any additional, baseline, or screening assessments (e.g., disposition of patients, demographics, medical history, site of infection and pathogen, and APACHE II score) will be summarized using descriptive statistics for continuous variables, or frequency counts and percentages for categorical variables.

Population Pharmacokinetics
A population PK analysis of plasma concentration-time data will be performed using non-linear mixed-effects modeling. Data from this trial may be combined with data from the Phase 1 trial in healthy adult volunteers and/or Phase 2 trial (STOP-AKI) in SA-AKI patients and included in an integrated PK analysis. The structural model will contain clearance and volume of distribution as fixed-effect parameters. The inter-patient variability in the parameter estimates and the random residual error in the data will be estimated with an appropriate model. Available patient characteristics will be tested as potential covariates affecting PK parameters. Details of the analysis will be given in a population PK analysis plan and the results of the population PK analysis will be presented in a separate report.

Interim Analyses
A maximum of four unblinded interim analyses will be conducted.
To maintain the blind and trial integrity, unblinded interim analysis outputs will be generated by a separate unblinded Biostatistics and Statistical Programming team not otherwise involved in the trial conduct. These unblinded interim results will exclusively be provided to the trial's DMC for review. The DMC will operate according to an approved DMC Charter and the DMC Chair will provide written recommendations on trial continuation or discontinuation to the Sponsor.
At the first interim analysis, recruitment may be stopped for futility. At subsequent interim analyses, recruitment may be stopped for futility or for early success (demonstration of superiority of recAP over placebo on 28 days mortality).

Success
The Lan-DeMets approximation of the O'Brien-Fleming alpha spending function will be used to determine the critical values for declaring success at interim and final analyses. Table 1 shows the nominal 1-sided significance levels at interim and final analyses for declaring success if data from exactly 700, 850, 1,000 and 1,400 patients in the main trial population will be available. In case the actual patient numbers differ, then the nominal significance levels will be re-calculated by the Lan-DeMets approach.

Table 1: Nominal 1-Sided Significance Levels for Success at Interim and Final Analyses (per the Lan-DeMets Approximation of the O'Brien-Fleming Alpha Spending Function)

Number of Patients Complete	Nominal One-Sided *p*-value
700	0.0015
850	0.0036
1,000	0.0067
1,400	0.0224

This trial will be considered a success if the 1-sided p-value from the primary efficacy analysis model for the treatment term is lower than the respective nominal 1-sided significance level.
Should the trial be stopped for success at an interim analysis, then the analysis of the secondary ednpoints within the sequrntial testing will proceed using the nominal 1-sided significance level as allocated to the primary endpoint at the time of the analysis.

Early futility
The predictive probability of success at the main trial population's maximum sample size (PPmax) of 1,400 patients will be used to determine if the trial should stop early for futility. This predictive probability calculation

combines the knowledge of the treatment effect observed in the trial with the uncertainty of the future data not yet observed. At the first interim analysis, the trial may stop for futility if the predictive probability of trial success is less than 15%. At each subsequent interim analysis, the trial may stop for futility if the predictive probability of trial success is less than 5%. Futility stopping in this trial is considered to be non-binding.