Signature Page

Name: DS-2020-022-5441
Title: A phase 2/3, randomized, open-label, parallel group, 3-arm, multicenter study investigating the efficacy and safety of intravenous administrations of emapalumab, an anti-interferon gamma (anti-IFNγ) monoclonal antibody, and anakinra, an interleukin-1(IL-1) receptor antagonist, versus standard of care, in reducing hyper-inflammation and respiratory distress in patients with SARS-CoV-2 infection.
Version: 3.0,CURRENT

Signed by:	Henrik Andersson (u100190)
Signed date:	2020-08-10 12:18:29
Role / Intention	Author Approval

Signed by:	Henrik Andersson (u100190)
Signed date:	2020-08-10 12:19:06
Role / Intention	Author Approval

Statistical Analysis Plan

Protocol number:	Sobi.IMMUNO-101
Title:	A phase 2/3, randomized, open-label, parallel group, 3-arm, multicenter study investigating the efficacy and safety of intravenous administrations of emapalumab, an anti-interferon gamma (anti-IFNγ) monoclonal antibody, and anakinra, an interleukin-1(IL-1) receptor antagonist, versus standard of care, in reducing hyper-inflammation and respiratory distress in patients with SARS-CoV-2 infection.
SAP Author:	Henrik Andersson
SAP version:	Final 3.0

Table of contents

1 Abbreviations and definition of terms

AE	Adverse event
ALT	Alanine aminotransferase
ANCOVA	Analysis of covariance
AST	Aspartate aminotransferase
Baseline	Before study drug administration at Visit 1 (Day 1)
CRP	C-reactive protein
CSR	Clinical study report
CT	Computed tomography
CXCL9	Chemokine (C-X-C Motif) ligand 9
DRC	Data review committee
ECMO	Extracorporeal membrane oxygenation
ECG	Electrocardiogram
FDA	Food and Drug Administration
FiO_2	Fraction of inspired oxygen
Hb	Hemoglobin
ICH	International council of harmonization of technical requirements for registration of pharmaceuticals for human use
IL	Interleukin
i.v.	Intravenous
IMV	Invasive mechanical ventilation
LDH	Lactate dehydrogenase
MedDRA	Medical Dictionary for Regulatory Activities
MEW	The modified early warning score
PaO_2	Partial pressure of oxygen
pCO_2	Carbon dioxide tension
pO_2	Oxygen tension
PT	Preferred term
RBC	Red blood cells

SAE	Serious adverse event
SAP	Statistical Analysis Plan
SARS-Cov-2	Severe acute respiratory syndrome coronavirus 2
SD	Standard deviation
Sobi	Swedish Orphan Biovitrum
SOC	System organ class
SpO_2	Peripheral capillary oxygen saturation
SRC	Safety review committee
TEAE	Treatment emergent AE
WBC	White blood cells
WHO	World Health Organization

2 Introduction

This SAP describes the planned analysis and reporting for the Sobi.IMMUNO-101 study, a phase 2/3, randomized, open-label, parallel group, 3-arm, multicenter study investigating the efficacy and safety of intravenous administrations of emapalumab and anakinra, versus standard of care, in reducing hyper-inflammation and respiratory distress in patients with SARS-CoV-2 infection.

The purpose of this SAP is to outline the planned analyses to be completed to support the Clinical study report (CSR) for protocol Sobi.IMMUNO-101. The planned analyses identified in this SAP will be included in potential regulatory submissions and/or future manuscripts. Also, exploratory analyses not necessarily identified in this SAP may be performed. Any post-hoc, or unplanned, analyses not identified in this SAP will be clearly identified in the CSR.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the FDA and ICH: Guidance on statistical principles in clinical trials (ICHE9). All work planned and reported for this SAP will follow internationally accepted guidelines, published by the American Statistical Association and the Royal Statistical Society, for statistical practice.

3 Study objectives and endpoints

3.1 Primary objective

The primary objective of this study is to assess the effect of emapalumab and anakinra on

hyperinflammation and pulmonary function in patients with SARS-CoV-2 infection.

3.2 Secondary objective

The secondary objective of this study is to evaluate the safety and tolerability profile, including evaluation of the immunogenicity of emapalumab and anakinra, of intravenous (i.v.) administrations of emapalumab and anakinra in patients with SARS-CoV-2 infection.

3.3 Exploratory objectives

An exploratory objective of this study is to evaluate the clinical status based on a 7-point ordinal scale.

An exploratory objective of this study is evaluate the PK of emapalumab and anakinra.

An exploratory objective of this study is to assess the effect of anakinra and emapalumab on CXCL9, IL-1, IL-6, sIL-2R and selected biomarkers relevant for hyperinflammation, whenever possible.

3.4 Study endpoints

3.4.1 Primary efficacy endpoint

The primary endpoint is treatment success, defined as not requiring any of the following by Day 15:

- Invasive mechanical ventilation (IMV), or
- Extracorporeal membrane oxygenation (ECMO).

3.4.2 Secondary endpoints supporting the primary objective

The secondary endpoints supporting the primary objective are:

- Time to mechanical ventilation.
- Change from baseline in MEWs score.
- Change from baseline in resting SpO_2.
- Change from baseline in PaO_2/FiO_2.
- Change from baseline in hemogasanalysis.
- Change from baseline in oxygen supplementation.
- Amelioration of the findings of high-resolution CT scan or X-ray of the chest.
- Change from baseline in hyperinflammatory parameters during treatment until Day 15 with measurements performed every 3 days:
    - Ferritin

  - LDH
  - D-dimers
- Change from baseline in other relevant laboratory parameters during treatment until Day 15 with measurements performed every 3 days:
  - WBC with differential counts
  - RBC
  - Hb
  - Platelet count
  - Fibrinogen
  - Complement C3/C4
  - Prothrombin time
  - Cardiac troponin
  - Liver tests (AST, ALT, total bilirubin levels)
  - CRP
  - Creatinine
  - Electrolytes (Sodium, Potassium, Calcium)
  - Glucose
- Overall survival
- Time to hospital discharge

3.4.3 Safety endpoints

Safety endpoints of this study are:

- Treatment-emergent serious adverse events (SAEs).
- Adverse events leading to premature discontinuation of study treatment.
- Infusion related reactions including anaphylactic/anaphylactoid reactions.
- Treatment emergent adverse events of special interest:
  - Emapalumab treatment group: All new infections.
  - Anakinra treatment group: Severe neutropenia.
- Treatment-emergent laboratory abnormalities.
- Presence of anti-drug antibodies (ADAs) against emapalumab and anakinra, and presence of neutralizing antibodies (Nabs) against anakinra.

3.4.4 Exploratory endpoints

Clinical status assessed by a 7-point ordinal scale. The scale consists of the following categories:

1. Death
2. Hospitalized, requiring invasive mechanical ventilation or extracorporeal membrane oxygenation (ECMO)
3. Hospitalized, requiring non-invasive ventilation or high flow oxygen devices
4. Hospitalized, requiring supplemental oxygen
5. Hospitalized, not requiring supplemental oxygen - requiring ongoing medical care
6. Hospitalized, not requiring supplemental oxygen - no longer requires ongoing medical care
7. Not hospitalized

PK parameters will be calculated for emapalumab and anakinra by non-compartmental analysis and population analysis, as applicable.

PD biomarkers: Change from baseline in CXCL9, IL-1, IL-6, sIL-2R and selected biomarkers relevant for hyperinflammation.

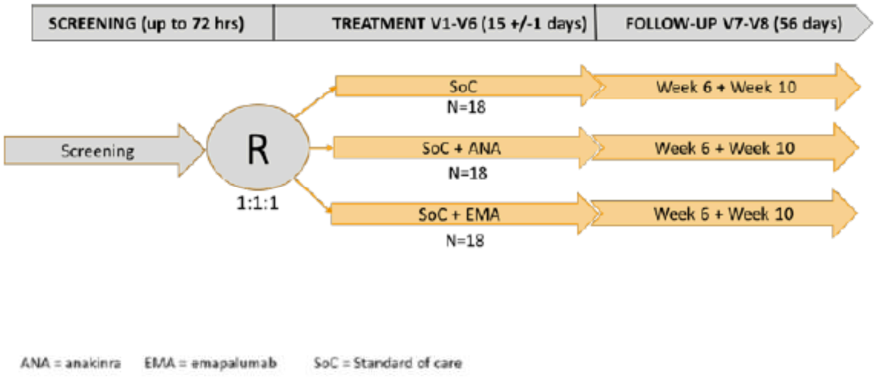
4 Study methods

4.1 Overall study design and plan

The study consists of screening, a 2-week treatment period and an 8-week follow-up period. The 2-week treatment period is open, and the patients will be randomized to treatment with emapalumab, anakinra or standard of care in a 1:1:1 ratio. Patients will be stratified based on glucocorticoid treatment at randomization to ensure an equal distribution between treatment arms. Emapalumab will be administered as i.v. infusions every 3rd day for a total of 5 infusions (Days 1, 4, 7, 10, 13). Anakinra will be administered as 4-times daily i.v. infusions for 15 days (Days 1 to 15). The primary endpoint will be evaluated at Day 15.

An early follow-up visit will be performed at Day 28. A follow-up by visit or phone call will be performed 4 and 8 weeks after the end of the treatment period (Weeks 6 and 10). The study duration for an individual patient will not exceed 10 weeks. The end of the study is defined as last patient, last follow-up visit/phone call. See Figure 1.

Figure 1 Study design



The study has a total sample size of 54 patients and consists of two stages. At the end of Stage 1 (interim analysis), the success rates are compared between each of the two treatment arms and standard of care and there is the potential to stop for futility or for efficacy of emapalumab or anakinra (or both).

4.2 Selection of study population

For inclusion into the study, patients should have documented presence of SARS-CoV-2 infection and be between the age > 18 to < 85 years at the time of screening. Further key requirements include the presence of respiratory distress and the presence of hyperinflammation, as defined in the protocol. Presence of certain medical conditions, as defined in the protocol, will preclude patients from entering the study.

4.3 Method of treatment assignment and randomization

This is a randomized, open-label study. The different treatment groups are;

Arm A: Emapalumab as add on to standard of care
Arm B: Anakinra as add on to standard of care
Arm C: Standard of care only

The sample size is 54 and the ratio between the treatment groups is 1:1:1, i.e., the same number of patients will be randomized to emapalumab, anakinra, or standard of care. Patients will be stratified based on glucocorticoid treatment at randomization to ensure an equal distribution between treatment arms. The randomization numbers are generated in blocks. Each block includes the three treatment groups per the ratio described above. An IWRS will be used for the randomization.

5 Sequence of planned analysis

5.1 Interim analyses

The study consists of one formal interim analysis (Stage 1) of the primary endpoint, a final analysis (Stage 2) of key primary and secondary endpoints, and an analysis after the final data cut of time-to-event endpoints including data from the 8-week follow-up period. The aim is to randomize equal numbers of patients into Stage 1 and into Stage 2 per treatment arm, but the Stage 1 interim analysis will be conducted when at least 9 patients have been accrued per treatment arm. There is the potential to stop for futility or for efficacy of emapalumab or anakinra (or both) at the end of Stage 1. The futility rule for stopping at the end of Stage 1 is binding.

More specifically:

- The emapalumab or anakinra arm (or both) will be stopped at the end of Stage 1 for futility if the one-sided p-value in favour of emapalumab/anakinra is > 0.690.
- The emapalumab or anakinra arm (or both) may be stopped for efficacy at the end of Stage 1 if the one-sided p-value in favour of emapalumab/anakinra is < 0.025.
- If the trial continues to Stage 2, efficacy will be declared at the end of Stage 2 if the one-sided p-value in favour of emapalumab/anakinra is < 0.159.

To exemplify certain operating characteristics of the design, Table 1 lists the outcomes that would lead to conclusions of efficacy at the ends of Stage 1 and Stage 2, or to a conclusion of futility at the end of Stage 1.

Table 1 Operating characteristics of treatment success rate in each group

Stage 1 Efficacy	**Emapalumab/anakinra**	**Standard of care**
The following outcomes (patients who are successful), and outcomes more extreme, give one-sided p < 0.025 for each pairwise comparison, to conclude efficacy at the end of Stage 1.	9 (100%)	4 (44%)
	8 (89%)	3 (33%)
	7 (78%)	1 (11%0
	6 (67%)	1 (11%)
	5 (56%)	0 (0%)
Stage 1 Futility		
The following outcomes, and outcomes more extreme, give one-sided p > 0.690 to conclude futility.	7 (78%)	9 (100%)
	6 (67%)	8 (89%)
	5 (56%)	7 (78%)
	4 (44%)	6 (67%)

	3 (33%)	6 (67%)
	2 (22%)	4 (44%)
	1 (11%)	3 (33%)
	0 (0%)	2 (22%)
Stage 2 Efficacy		
The following outcomes (patients who are successful), and outcomes more extreme, give one-sided $p < 0.159$ for each pairwise comparison, to conclude efficacy at the end of Stage 2.	18 (100%)	15 (83%)
	17 (94%)	13 (72%)
	16 (89%)	12 (67%)
	15 (83%)	11 (61%)
	14 (78%)	10 (56%)
	13 (72%)	9 (50%)
	12 (67%)	8 (44%)
	11 (61%)	7 (39%)
	10 (56%)	6 (33%)
	9 (50%)	5 (28%)
	8 (44%)	4 (22%)
	7 (39%)	3 (17%)
	6 (33%)	2 (11%)
	5 (28%)	1 (6%)
	4 (22%)	0 (0%)
	3 (17%)	0 (0%)

The calculations on the operating characteristics of this design have been undertaken using PASS, Version 14: Group-Sequential Tests for Two Proportions (Simulation).

The Data review committee (DRC) composed of experts in intensive care, inflammation, infectious diseases will be involved in study oversight and interpretation of the study results.

Clinical safety information (AEs, SAEs, laboratory values, ECGs, vital signs, and study-specific labs/examinations, as required) will be monitored and reviewed on a continuous basis by the Safety Review Committee (SRC) by periodically monitoring clinical studies activities from protocol conception to database closure. There will be separate written operating procedures for

the SRC.

5.2 Analyses and reporting

The analyses identified in the SAP will be performed at different stages of the study. The primary endpoint will be analyzed at Stage 1 and Stage 2. Most of the other endpoints will be analyzed at Stage 2 after all patients have completed Day 15, and the remaining ones that are assessed during the follow-up period will be analyzed at the end of the follow up, as specified in Section 12.2.1. The SAP will be finalized, locked and signed prior to the Stage 1 interim analysis.

Any post-hoc analyses included in the CSR, which were not identified in this SAP, will be clearly identified as such in the relevant section of the CSR.

6 Sample size determination

The study will enroll a total of 54 patients, 18 per arm according to the 1:1:1 randomization.

The sample size has been estimated based on the following assumptions:

- An overall one-sided significance level for efficacy of 0.097 (9.7%) and a power of 74% for each comparison under the assumption that the true success rates are 50% in the SoC group increasing to 80% in the emapalumab or anakinra groups.
- The study will consist of two stages. The Stage 1 interim analysis will be conducted after the accrual of at least 9 patients per treatment arm.
- There is the potential to stop for futility or for efficacy of emapalumab or anakinra (or both) at the end of Stage 1, but only the futility rule is binding and irrespective of the Stage 1 results, the study will not be stopped for efficacy but continue to accrue efficacy and safety data.

The value for the type I error has been chosen in recognition of the urgent unmet medical need to allow the identification of a signal, at least, from a statistical perspective. Having seen a statistical signal, it is then a matter of evaluating whether the observed treatment differences represent clinically relevant effects that can satisfy that unmet need. Assuming no untoward safety signal is observed and no futility stopping occurs at the Stage 1 interim, the decision may be taken to increase the sample size beyond 54 patients in order to accrue further safety and efficacy data. Details of this would be described in a protocol amendment. In case the outcome is statistically convincing, i.e. statistically significant on a 5% level in the stage 2 analysis, efficacy will be considered confirmed and the results may be used to seek regulatory approvals. The DRC will provide support in such discussions.

7 Analysis populations

The following analysis sets will be used in the statistical analyses.

7.1 Modified intention-to-treat population

All efficacy analyses will be conducted on the modified intention-to-treat (mITT) population which will comprise all randomized patients except patients who did not receive study treatment or patients who tested negative to SARS-Cov-2 diagnosis confirmation by PCR testing.
In case of exclusion due to these criteria, additional patients may be recruited to ensure an adequate number of evaluable patients for the efficacy analysis.

Patients will be included in the group to which they were randomized for all efficacy evaluations.

7.2 Safety population

The Safety population will comprise all patients who received at least one dose of study treatment or standard of care.

Patients will be included according to treatment received for all evaluations of safety.

8 General issues for statistical analysis

Results will be presented as the estimated value for each treatment group, emapalumab, anakinra and standard of care, the estimated difference between emapalumab/anakinra and standard of care, the associated confidence interval and p-value. P-values from statistical analyses will be presented to three decimal places with values below 0.001 displayed as <0.001 and confidence intervals will be presented to one more decimal place than the raw data.

Continuous data will be summarized using descriptive statistics: n, mean, SD, median, minimum and maximum, unless otherwise indicated. Minimum and maximum will be presented to the same number of decimal places as the raw data and mean, SD, and median will be presented to one more decimal place than the raw data.

Categorical data will be summarized using counts and percentages. Percentages will be suppressed when the count is zero, however the category will still be displayed. The denominator for all percentages will be the number of patients within the treatment group for the population of interest, unless otherwise indicated. Percentages will be presented to one decimal place.

All efficacy endpoints will be analyzed and summarized for the mITT population and safety data will be presented for the safety population.

Statistical analyses will be performed using SAS software Version 9.4 (SAS Institute Inc, Cary, North Carolina, United States).

8.1 Time windows

For the purpose of the statistical analysis and if not otherwise specified, all the variables (primary, secondary and exploratory variables) will be assigned to the visit/assessment timepoint in which they were collected depending on the following analysis time windows in Table 1.

Table 2 Time windows

Visit	Study Day	Study Day window
1	1 (Baseline)	≤1
2	4	2-5
3	7	6-8
4	10	9-11
5	13	12-14
6	15	15-16
7	28	Early follow-up
7/TC	Week 6	Only follow-up of survival and hospital discharge and continued AE reporting in accordance with the protocol.
8/TC	Week 10	

Unscheduled visit data will be considered when assigning assessments to the analysis visit. If more than one assessment of the variable (scheduled or unscheduled) falls in the same time window but on different days, the closest to the scheduled visit day will be taken, or the earlier, in the event the values are equidistant from the nominal visit date. If several measurements are collected during the selected day, the average (for numeric values)/worst (for categorical values) of the measurements will be taken.

In summaries of extreme values as opposed to visit-based summaries, all post baseline values collected are used including those collected at unscheduled visits regardless if the value is closest to the scheduled visit date or not. For summaries at a patient level, all values should be included, regardless of whether they appear in a corresponding visit-based summary, when deriving a patient level statistic such as a maximum.

Listings will display all values contributing to a time point for a patient.

8.2 Handling of missing data and outliers

8.2.1 Imputation of efficacy endpoints

For the primary endpoint (treatment success by Day 15), patients who discontinue study treatment (or discontinue the study if randomized to standard of care) prior to Day 15 and had IMV or ECMO on their last assessment by the time of study treatment discontinuation, will be considered not having achieved treatment success by Day 15. Patients discontinuing prior to Day 15 without ever having initiated IMV/ECMO, will be considered having achieved treatment success by Day 15. Patients discontinuing prior to Day 15 due to having significantly improved will be considered having achieved treatment success. Patients who died by Day 15, have not achieved treatment success regardless of IMV/ECMO use by Day 15.

For statistical analyses of change-from-baseline efficacy endpoints at Day 15, the last observation carried forward (LOCF) approach will be used to impute missing data, using the last post-baseline observation available for each patient.

For other efficacy presentations no imputation will be performed, i.e., all presentations will be based on observed data.

8.2.2 Imputation of adverse event start dates

To assess if AEs are treatment emergent, missing onset dates will be imputed according to the following rules:

- If day is missing and month and year of AE onset is the same as the month and year of first dose of study treatment (or date of randomization for patients on standard of care), then the onset date of the AE will be assumed to be the same as the date of first dose of study treatment (or date of randomization for patients on standard of care).
- If day and month is missing and year of AE onset is the same as year of first dose, then the date of onset will be assumed to be the same as the first dose of study treatment (or date of randomization for patients on standard of care).
- Otherwise, if day is missing, then use YYYY-MM-01, and if day and month are missing, then use YYYY-JAN-01.
- Completely missing onset dates will not be imputed. AEs with a completely missing onset date will be considered treatment-emergent.

8.2.3 Imputation of previous/concomitant medication start/stop dates

To assess if medications are prior or concomitant, missing onset dates will be imputed according to the following rules:

- If day is missing and month and year are both available, use 01 of the month or the date of first dose of study treatment (or date of randomization for patients on standard of care) whichever is later.

- If only the year is available, use 01 Jan of the year or the date of first dose whichever is later.
- Completely missing onset dates will not be imputed.

Likewise, to assess if a medication is prior, missing stop dates will be imputed according to

- If the day is missing: Assume the last day of the month.
- If the month is missing: Assume 31 Dec of the year.

Completely missing stop dates will not be imputed.

8.3 Multicenter studies

No effects of centers will be evaluated in this study due to the small numbers at individual sites.

8.4 Multiple comparisons and multiplicity

There will be no adjustments for multiplicity for the two pairwise treatment comparisons for the primary endpoint. Multiplicity for the sequential comparisons at the end of Stage 1 and at the end of Stage 2 are however accounted for by the design.

8.5 Derived and computed variables

8.5.1 Treatment success

Treatment success is achieved for patients not requiring IMV or ECMO by Day 15. If IMV or ECMO is required at any visits prior to Day 15/date of early discontinuation, but no longer required at Day 15/date of early discontinuation, treatment success has been achieved. If IMV or ECMO is not initiated during the study, treatment success has been achieved. As stated in Section 8.2.1, patients who discontinue study treatment (or discontinue the study if randomized to standard of care) prior to Day 15 and had ECMO or IMV on their last assessment by the time of study treatment discontinuation, will be considered not having achieved treatment success by Day 15. Patients who discontinue prior to Day 15 without ever having initiated IMV/ECMO, have achieved treatment success by Day 15. Patients discontinuing prior to Day 15 due to having significantly improved will be considered having achieved treatment success. Patients who died by Day 15, have not achieved treatment success regardless of IMV/ECMO use by Day 15.

8.5.2 Time to mechanical ventilation, hospital discharge and overall survival

Time to mechanical ventilation is defined as the time from randomization until mechanical ventilation (i.e. date of mechanical ventilation or censoring – date of randomization + 1). Any patient not known to have initiated mechanical ventilation at the time of analysis will be censored based on the last recorded date on which the patient was known to not requiring

mechanical ventilation (e.g. date of Day 15 visit, date of early discontinuation, or date of death). A separate analysis of time to IMV *or* ECMO will also be conducted with the same definition as above, but both IMV and ECMO will be considered as an event.

Overall survival is defined as the time from randomization until death (i.e. date of death or censoring – date of randomization + 1). Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive.

Time to hospital discharge is defined as the time from randomization until hospital discharge (i.e. date of hospital discharge or censoring – date of randomization + 1). Any patient not known to have been discharged at the time of analysis will be censored based on the last recorded date on which the patient was known to be hospitalized (e.g. date of visit, date of early discontinuation, or date of death).

8.5.3 MEWs score

Modified early warning system score (MEW) will be assessed at Visit 1 (Day 1), Visit 6 (Day 15) and Visit 7 (Day 28). The total score is a sum of the five categorized components that will be assigned a score between 0 to 2 or 0 to 3. A higher score is worse. The components are systolic blood pressure, heart rate, respiratory rate, temperature and the alert/voice/pain/unresponsive score.

8.5.4 SpO_2

Resting peripheral capillary oxygen saturation (SpO_2) will be measured 3 times per day. An average value across the three assessments will be calculated before the analysis of change from baseline.

8.5.5 PaO_2/FiO_2

The ratio of partial pressure of oxygen (PaO_2) and fraction of inspired oxygen (FiO_2) will be calculated.

9 Patient disposition

The number and percentage of patients who were enrolled, randomized, who initiated and who completed or discontinued treatment early, together with the reasons for discontinuations, will be presented by treatment group. The number of patients in the All Treated population will be presented.

10 Demographics and baseline characteristics

10.1 Demographics and baseline characteristics

Baseline characteristics including weight, height and oxygen levels, demographic data including age, sex, race, and ethnicity, and physical examination data, will be presented by descriptive statistics.

10.2 ECG and vital signs at screening

ECG and vital signs collected at screening will be summarized with descriptive statistics.

10.3 Medical history

The number and percentage of patients with medical history of particular interest will be presented separately. Other medical history will be summarized by system organ class (SOC) and preferred term (PT).

Medical history data will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA).

11 Prior and concomitant medication

All prior and concomitant medication will be coded using the latest version of World Health Organization drug dictionary (WHO Drug).

The preferred term grouped by ATC level 4 will be used for presentation and sorted in descending order of frequency in total (the treatment groups together) at the top in the table. Prior medication, concomitant medication at randomization and onset of new concomitant medication after randomization during the study will be summarized separately.

12 Efficacy analyses

12.1 Primary efficacy endpoint

For the analysis of the primary endpoint, each of the two pairwise comparisons, emapalumab/anakinra versus standard of care, at both the end of Stage 1 and at the end of Stage 2, will be undertaken using Fisher's exact test comparing the proportion of patients with treatment success.

The null and alternative hypotheses with respect to the primary endpoint are defined as:

H_0: $P_{treatment} = P_{SoC}$

H_A: $P_{treatment} > P_{SoC}$

where P is the proportion of patients with treatment success as defined for the primary endpoint. The one-sided p-value from Fisher's exact test will be supplemented by the presentation of one-sided exact [Santner et al 1980] confidence intervals of the differences in proportion between emapalumab/anakinra and standard of care, with confidence coefficients of 97.5% at the end of Stage 1 and 84% at the end of Stage 2. Two-sided confidence intervals of the proportions in each group will also be presented, using the method of Clopper-Pearson [Clopper et al 1934].

The primary analysis will not adjust for the stratification factor glucocorticoid use at inclusion in order to avoid incorrect weights being attributed to any classes of the factors with few data points, particularly at Stage 1. However, a sensitivity analysis of the primary endpoint using exact logistic regression will be performed at Stage 1 and 2. This will allow for conditional testing of the treatment effect given the stratification factor. The results of the analysis will be presented in terms of model-adjusted treatment success rates for each treatment group, and p-values and confidence intervals for the odds ratio of the pairwise comparisons to the standard-of-care arm.

The individual success rate profiles will be displayed in a Swimmer plot with bars representing the time in the study and markers for whether IMV/ECMO was used at each visit. The purpose is to show onset of IMV/ECMO and how it evolves over time in relation to treatment.

The relationship between achieving treatment success/no treatment success and the change in hyperinflammatory parameters will be explored by treatment group in Box-Whisker plots with individual scatter.

12.2 Secondary endpoints supporting the primary objective

The secondary endpoints supporting the primary objective will not be analyzed at Stage 1.

12.2.1 Time to mechanical ventilation, hospital discharge and overall survival

Analysis of the time to mechanical ventilation from the point of randomization will be undertaken by plotting Kaplan-Meier curves for each of the 3 treatment groups and by pairwise comparisons (emapalumab/anakinra versus standard of care) using the log-rank test. Hazard ratios will be estimated using the Cox proportional hazards model and these will be presented together with 90% two-sided confidence intervals. Time to IMV *or* ECMO will be analyzed using the same approach.

Overall survival (time to death from the point of randomization) and time to hospital discharge (from the point of randomization) will not be analyzed at the end of Stage 2, but at the end of the follow-up period. This endpoint will be analyzed as for time to mechanical ventilation.

12.2.2 Change from baseline in MEWs score

Change from baseline in MEWs score at Day 15/LOCF will be analyzed using analysis of covariance (ANCOVA) including treatment arm as a fixed factor and baseline MEWs score as a covariate. Least square mean changes per group, associated 90% two-sided CI, and p-values for the comparison vs. standard of care, will be presented.

12.2.3 Change from baseline in resting SpO2 and oxygen supplementation

Change from baseline in resting SpO_2 result and in amount of oxygen supplementation during treatment over time will be summarized using descriptive statistics (mean, SD, median, Q1, Q3, min and max) for each of these parameters. In addition to the change from baseline, absolute values at baseline and at each visit will be summarized.

Furthermore, change from baseline in resting SpO2 and oxygen supplementation at Day 15/LOCF will be analyzed using ANCOVA as described in 12.2.2.

12.2.4 Change from baseline in PaO_2/FiO_2

Change from baseline (screening) during treatment by visit will be summarized using descriptive statistics as described in 12.2.3.

Furthermore, change from baseline in PaO2/FiO2 at Day 15/LOCF will be analyzed using ANCOVA as described in 12.2.2.

12.2.5 Change from baseline in hemogasanalysis

Change from baseline in the hemogasanalysis parameters (pH, carbon dioxide tension (pCO_2), oxygen tension (pO_2), electrolytes, lactate and hemoglobin) during treatment by visit will be summarized using descriptive statistics as described in 12.2.3.

Furthermore, change from baseline in hemogasanalysis parameters at Day 15/LOCF will be analyzed using ANCOVA as described in 12.2.2.

12.2.6 Amelioration of the findings of high-resolution CT scan or X-ray of the chest

Amelioration in the findings of high-resolution CT scan or X-ray of the chest will be summarized in a shift table showing the number and percentage of patients that had normal, abnormal (not clinically significant) or abnormal (clinically significant) findings at screening and at Day 15.

12.2.7 Change from baseline in hyperinflammatory parameters

Change from baseline in hyperinflammatory parameters (ferritin, LDH and D-dimers) during treatment by visit until Day 15 will be summarized using descriptive statistics as described in 12.2.3.

Furthermore, change from baseline in hyperinflammatory parameters at Day 15/LOCF will be analyzed using ANCOVA as described in 12.2.2.

In addition, for each hyperinflammatory parameter, a summary will be included with the number and percentage of patients that had:

- at least 50% improvement from baseline to Day 15, *and*
- normalization by Day 15

12.2.8 Change from baseline in other relevant laboratory parameters

Change from baseline in other relevant laboratory parameters during treatment by visit until Day 15 will be summarized using descriptive statistics as described in 12.2.3.

Furthermore, change from baseline in other relevant laboratory parameters at Day 15/LOCF will be analyzed using ANCOVA as described in 12.2.2.

12.3 Subgroup Analyses

No formal subgroup analyses are planned.

12.4 Exploratory endpoints

12.4.1 Clinical status assessed by a 7-point ordinal scale

Clinical status using the 7-point ordinal scale is an exploratory endpoint and the data collected will be summarized to show the number and percentage of patients in each category over time.

12.4.2 Pharmacokinetics

PK data will be listed and summarized at each visit with descriptive statistics.

12.4.3 Pharmacodynamics

Change from baseline in CXCL9, IL-1, IL-6, sIL-2R and any other selected biomarkers relevant for hyperinflammation that are collected, will be summarized with descriptive statistics by visit.

13 Safety analyses

In addition to the summaries of safety data for the CSR presented in subsequent sections, listings and tabulations may be provided ad-hoc for the SRC to support their review, as appropriate, in addition to individual case reports etc. from the study centres. Further details will be provided in the SRC charter.

13.1 Adverse events

All AEs will be coded using MedDRA version 21.1.

A treatment-emergent AE (TEAE) is defined as an AE with a start date/time after randomization.

The number and percentage of patients with at least one TEAE, at least one serious TEAE, including death, at least one non-serious TEAE, any related TEAE, any fatal TEAE, any TEAE leading to study treatment withdrawn and any TEAE leading to study withdrawal will be summarized.

The number and percentage of patients with at least one TEAE will be summarized in frequency tables by treatment, system organ class and preferred term. Separate tables will present TEAEs classified by relationship to study treatment or to standard of care and TEAEs classified by maximum severity. These tables will be presented by PT in descending order of frequency across all groups. Percentages will be based on the number of patients in the All Treated population for the specific treatment group.

13.1.1 Serious adverse events

The number and percentage of patients with at least one serious TEAE will be tabulated by SOC and PT. SAEs will also be listed. The listing will include information on treatment group, age, sex, race, start/stop date of event, SOC/PT, causality, severity, action taken, and outcome.

13.1.2 Adverse events leading to withdrawal

The number and percentage of patients with TEAEs leading to study withdrawal will be presented by SOC and PT.

13.1.3 Adverse events of special interest

The number and percentage of patients with treatment emergent adverse events of special interest will be presented:

- Emapalumab treatment group: Infections caused by pathogens potentially favored by IFN-γ neutralizations such as mycobacteria, salmonella, shigella, herpes zoster and histoplasma capsulatum, and severe infusion-related reactions.
- Anakinra treatment group: Severe neutropenia.

13.1.4 Deaths

Details of any deaths will be listed together with SAEs as detailed in 13.1.1.

13.2 Drug exposure

Exposure to emapalumab, anakinra and methylprednisolone background therapy, respectively, will be summarized with descriptive statistics.

13.3 Laboratory data

The laboratory data evaluation will be based on the results from the local laboratories. For potential queries of laboratory data outside of reference ranges, ±10 percent will be allowed to account for variability of certain lab parameters in the different local laboratories.

The laboratory safety data will be presented as actual values and change from baseline values over time by descriptive statistics. The number and percentage of patients with low, normal, or high laboratory values at baseline versus subsequent visits will be presented using shift tables.

13.4 Vital signs

Vital signs will be presented as absolute and change from baseline values by visit with descriptive statistics.

13.5 ECG

ECG will be presented as absolute and change from baseline values by visit with descriptive statistics.

14 References

Clopper C et al. The use of confidence or fiducial limits illustrated in the case of the Binomial. Biometrika 1934;26: 404-413.

Santner, T. J., and Snell, M. K. (1980). Small-Sample Confidence Intervals for and in Contingency Tables. Journal of the American Statistical Association. 1980;75:386–394.