Statistical Analysis Plan (SAP)

Protocol Title:	A Phase 2/3, Open-label, Single Arm, Multicenter Study to Evaluate the Efficacy, Safety, and Pharmacokinetics of Emapalumab in Adult Patients with Hemophagocytic Lymphohistiocytosis
Protocol Version No./Date:	3.0/09-MAR-2020
CRF Version No./Date:	2.0/30-JUL-2020
SAP Version No./Date:	1.0/24-JUN-2021

1.0 Approvals

Sponsor	
Sponsor Name:	Swedish Orphan Biovitrum AG
Representative/Title:	Anna Stoltenberg/Statistical Science Director
Signature/Date:	DocuSigned by: Anna Stoltenberg 6/29/2021
Representative/Title:	Josefin Blomkvist/Statistical Scientist
Signature/Date:	DocuSigned by: Josefin Blomkvist 6/29/2021
PRA	
Biostatistician/Title:	Sophie Collingborn/Principal Biostatistician
Signature/Date:	Signed by: CollingbornSophie Principal Biostatistician Reason: I agree to the terms defined by the placement of my signature on this document. Date & Time: 30 Jun 2021 10:43 AM +02:00

(NOTE: Electronic signatures should only be used if all parties have the ability to eSign.)

2.0 Table of Contents

3.0 Introduction

This statistical analysis plan (SAP) describes the planned analysis and reporting for Sobi Protocol NI-0501-10 titled, A Phase 2/3, Open-label, Single Arm, Multicenter Study to Evaluate the Efficacy, Safety, and Pharmacokinetics of Emapalumab in Adult Patients with Hemophagocytic Lymphohistiocytosis, dated March 9, 2020.

The study was planned to enroll adult patients diagnosed with Hemophagocytic Lymphohistiocytosis (HLH) in two different cohorts based on disease etiologies; specifically, newly diagnosed patients with malignancy-associated HLH (M-HLH), and newly diagnosed or previously treated patients with non-malignancy-associated HLH.

The study planned to include an initial phase enrolling 10 patients. If no untoward significant safety signal attributable to emapalumab treatment, as judged by an independent Data Monitoring Committee (iDMC) is identified, enrollment was to continue. In addition to the safety data that is being monitored by the iDMC, adaptive assessments for futility and early efficacy will be performed by the iDMC. Details of the analyses for futility and early efficacy assessments performed by the iDMC are included in this SAP. All other details of assessments performed by the iDMC are provided in the iDMC Charter and iDMC SAP (quarterly iDMC review/meeting).

The planned statistical methods and analyses have been selected taking into account, where applicable, relevant sections of the International Council on Harmonization (ICH) E9 Guideline “Statistical Principles for Clinical Trials,” and the ICH E3 Guideline “Structure and Content of Clinical Study Reports.” The analyses identified in this SAP may be used for generation of the clinical study report (CSR), regulatory submissions, or future manuscripts, as appropriate. Any exploratory analyses not necessarily identified in this SAP may be performed to further examine study data; such analyses will be clearly identified as such in the final CSR.

The SAP was planned to be finalized prior to the first adaptive assessment for futility and early efficacy, i.e., before the 10 patients with M-HLH have completed treatment. However, the study was prematurely discontinued prior finalization of the SAP. The analysis presented in the SAP are designed to take account of early discontinuation. This SAP should be read in conjunction with the study protocol, version 3.0 and the case report form (CRF), version 2.0. Any further changes to the protocol or CRF may necessitate updates to the SAP.

3.1 Changes from Protocol

The following changes from the protocol have been added in this SAP:

- The Per Protocol Analysis Set was not used.
- Since the study was prematurely discontinued, the number of TFLs were reduced considering the low number of subjects. For example duration of response will not be summarized using Kaplan-Meier (KM) methods.

4.0 Study Objectives

The primary objective is to assess the efficacy of emapalumab in adult patients with HLH, as measured by Overall Response Rate (ORR).

The secondary objectives are as follows:

- To assess the efficacy of emapalumab as measured by Overall Survival, time to response, best response on treatment, duration of response, and other efficacy parameters.
- To evaluate the safety and tolerability of emapalumab.
- To determine the pharmacokinetic (PK) profile of emapalumab.
- To determine the pharmacodynamic (PD) profile of emapalumab.

5.0 Study Design

Study NI-0501-10 is an open-label, single arm, multicenter, Phase 2/3 interventional study.

The study was planned to enroll adult patients diagnosed with HLH in two different cohorts based on disease etiologies; specifically, newly diagnosed patients with malignancy-associated HLH (M-HLH), and newly diagnosed or previously treated patients with non-malignancy-associated HLH.

The study includes an initial phase enrolling 10 patients in a cohort to perform an adaptive assessment for futility and early efficacy for the independent Data Monitoring Committee (iDMC) to review. If no untoward, significant safety signal, attributable to emapalumab treatment, as judged by the iDMC, is identified, enrollment will continue. Additionally, reviews by the iDMC will take place after every subsequent 5 patients per cohort, i.e. 15, and 20, for assessments of early efficacy and futility. The iDMC will also meet on a quarterly basis to confirm that the safety profile of emapalumab remains favorable and the benefit/risk of treatment remains positive. The details of the iDMC are provided in the iDMC Charter and iDMC SAP.

The patient's participation in the study comprises three parts: screening, treatment period, and follow-up. After giving written informed consent, patients who are willing to participate in the study will undergo screening assessments within 2 weeks prior to the baseline visit. Patients must fulfill all the entry criteria for participation in the study and the results must be available prior to the baseline visit.

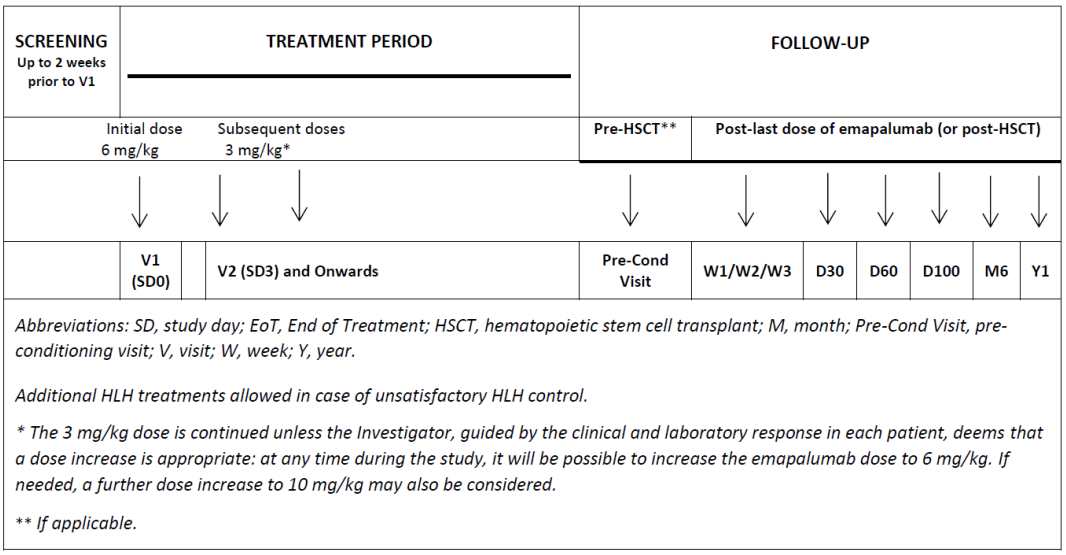
Eligible patients will be administered emapalumab by intravenous (IV) infusion over a period of 1 to 2 hours, depending on the volume to be infused, at an initial dose of 6 mg/kg. Emapalumab treatment will be continued at a dose of 3 mg/kg, every 3 days for the first two weeks, and then twice-a-week, until a clinically satisfactory response is achieved. If HSCT is planned, treatment can be administered until HSCT. Patients will undergo a Week 4 Assessment Visit 3 days after the last emapalumab infusion administered in Week 4. Treatment may be shorter, and in such instance an End of Treatment (EOT) Visit will be performed 3 days after the last administration of emapalumab. Whenever a patient receives treatment beyond Week 4, both a Week 4 Assessment Visit and an EOT Visit will be performed.

After treatment discontinuation (for any reason), patients will continue in the study for long-term follow-up until 1 year after last emapalumab administration (or after HSCT, if performed). If HLH recurs or reactivates during the off-drug follow-up period, treatment with emapalumab can be reinstated, upon discussion with the Sponsor. Patients will enter the follow-up period after completion of the re-treatment.

There will be a 1-Year Visit at the end of the follow-up period which represents the last visit of the study. The assessments planned for the 1-Year Visit will also be performed in patients who prematurely withdraw from the study. Patients will not be followed for any reason after consent has been withdrawn.

A schematic of the study design is shown in Figure 6-1.

Figure 6-1 NI-0501-10 Study Design



Abbreviations: SD, study day; EoT, End of Treatment; HSCT, hematopoietic stem cell transplant; M, month; Pre-Cond Visit, pre-conditioning visit; V, visit; W, week; Y, year.

Additional HLH treatments allowed in case of unsatisfactory HLH control.

** The 3 mg/kg dose is continued unless the Investigator, guided by the clinical and laboratory response in each patient, deems that a dose increase is appropriate: at any time during the study, it will be possible to increase the emapalumab dose to 6 mg/kg. If needed, a further dose increase to 10 mg/kg may also be considered.*

*** If applicable.*

5.1 Sample Size Considerations

The primary analysis was planned to evaluate the ORR at Week 4 (or EOT Visit in case of earlier treatment discontinuation) per etiology using an exact binomial test with a one-sided significance level of 5% to calculate the p-value.

According to the design of this study, the initial sample size planned for the study is 10 patients per etiology. The sample size may be expanded up to 25 patients for a particular etiology after iDMC review (see details in Section 9.0). The study was terminated after 7 patients had been recruited.

5.2 Randomization

This is a single arm study; therefore, randomization is not applicable.

6.0 Study Endpoints and Analysis Sets

6.1 Study Endpoints

6.1.1 **Efficacy Endpoints**

Any analysis related to primary and secondary efficacy endpoint will not be performed since the number of patients is very low and no conclusion can be reached.

The primary efficacy endpoint of this study is:

- Overall Response, i.e., achievement of either a Complete Response (CR) or Partial Response (PR) at Week 4 (or EOT if earlier). Criteria for the definition of Overall Response are provided in Table 7-1. A more detailed specification for the implementation of these criteria in the derivation of Overall Response is presented in Appendix 12.1.

The secondary efficacy endpoints that will be assessed in this study are as follows:

- Best response on treatment
- Overall Response at EOT
- Overall Survival (OS)
- Time to Complete or Partial Response
- Duration of response
- HLH Relapse

Table 7-1 Definition of Overall Response

Overall Response	
Complete Response	Complete response is adjudicated if: • No fever = body temperature <37.5°C • Normal spleen size • No cytopenia* = Absolute Neutrophil Counts ≥1.0x10^9/L and platelet count ≥100x10^9/L [absence of Granulocyte-colony-stimulating factor (G-CSF) and transfusion support must be documented for at least 4 days to report no cytopenia] • No hyperferritinemia = serum level is <2000 µg/L • No evidence of coagulopathy, i.e., normal D-Dimer and/or normal (>150 mg/dL) fibrinogen levels • No neurological and CSF abnormalities attributed to HLH

	• No sustained worsening of sCD25 (as indicated by at least two consecutive measurements that are >2-fold higher than baseline)
Partial Response	Partial Response is adjudicated if: • Improvement (>50% change from baseline or normalization) of at least 3 HLH clinical and laboratory abnormalities (including Central Nervous System (CNS) abnormalities)

* Not to be assessed if not HLH-related, such as cytopenia induced by drugs administered to treat the underlying disease (e.g. chemotherapy)

6.1.2 Safety Endpoints

Safety and tolerability of emapalumab will be assessed as follows:

- Incidence, severity, causality and outcomes of adverse event (AEs) (serious and non-serious)
- Clinically significant changes in blood tests, physical examinations, vital signs, ECGs.

6.1.3 Pharmacokinetic and Pharmacodynamic Endpoints

Since the study was prematurely discontinued, the PK/PD data will be only be listed in the following listings for the All Treated Analysis Set:

- PD measurements for:
  - Levels of circulating total IFNγ (free + bound)
  - Markers of IFNγ neutralization, including CXCL9
  - Other relevant markers, including but not limited to soluble CD25 (sCD25) and interleukin 6 (IL-6)
- PK measurements which will present serum concentrations of emapalumab at specified timepoints
- Immunogenicity measurements which will present the anti-drug antibodies (ADAs) against emapalumab.

A separate analysis plan and report will not be created for the PK/PD analyses and characterization.

6.2 Analysis Sets

6.2.1 All Treated Analysis Set

The All Treated Analysis Set will include all patients entering the study who receive any part of an infusion of study drug (e.g. patients who receive at least one dose or at least one incomplete dose).

6.3 Protocol Deviations

The study specific Protocol Deviation Guidance Document defines all protocol deviations considering operational aspect of the study conduct, including deviations as a result of the COVID-19 pandemic. Protocol deviations comprise the following categories:

- Inclusion criteria
- Exclusion criteria
- Study Drug
- Safety Assessment
- Lab/Endpoint Data
- Visit Window

- Informed Consent
- Other

Per PRA processes, all protocol deviations data will be entered into the system of record (PSO).

The study team and the Sponsor will conduct ongoing reviews of the deviation data from PSO throughout the study and prior to the database lock and prior to each adaptive assessment of Overall Response, adjusting the deviation criteria as appropriate.

- The following protocol deviations are considered to significantly impact the statistical analysis outcomes and will be classified as Major: Violation of in- or exclusion criteria:
    - Violation of inclusion criterion #2: Fulfillment of HLH-2004 clinical criteria, i.e. at least five out of eight criteria met.
    - Violation of inclusion criterion #3: Patients diagnosed with M-HLH must be treatment naïve.
    - Violation of inclusion criterion #5: JAK inhibitors, if taken, must be discontinued before emapalumab initiation.
    - Violation of exclusion criterion #1: Patient has primary HLH.
    - Violation of exclusion criterion #2: Current or scheduled administration of therapies known to trigger the cytokine release syndrome.
    - Violation of exclusion criterion #3: Current or scheduled administration of PD-1/PD-L1/ CTLA-4 inhibitors.
    - Violation of exclusion criterion #7: Leishmania infections.
- Concurrent treatment with a non-permitted drug

7.0 Conventions, Derivations, and Data Handling

7.1 Definitions and Derivations

7.1.1 **Study Day**

Throughout this SAP, study days are defined as follows:

- Study Day 1 is the day of first administration of IMP.
- For study days on or after the first administration of IMP, study day will be calculated as:

$$Study\ Day = Date\ of\ day - Date\ of\ first\ administration\ of\ IMP + 1\ .$$

- For study days before the first administration of IMP, study day will be calculated as:

$$Study\ Day = Date\ of\ day - Date\ of\ first\ administration\ of\ IMP.$$

Note: In the schedule of events in the clinical study protocol, the day of first IMP administration is defined as SD0, which is not a valid study day in Clinical Data Interchange Standards Consortium (CDISC). Consequently, study day numbering in this SAP and for statistical programming will use the above definitions rather than the study days defined in the clinical study protocol.

7.1.2 **Durations**

In general, duration of events (days) will be calculated as

$$Duration\ (days) = End\ date\ of\ event - Start\ date\ of\ event + 1$$

7.1.3 **Baseline**

Unless otherwise specified, baseline is defined as the last non-missing value before the start of first administration of IMP, including the pre-infusion value on SD1 or screening value if necessary.

7.1.4 Change from Baseline

Change from baseline will be defined only for post-baseline timepoints as:

$$Change\ from\ Baseline = Post\ baseline\ value - Baseline\ value.$$

7.1.5 Fold Change from Baseline

Fold change from baseline will be defined for post-baseline timepoints as

$$Fold\ Change\ from\ Baseline = \frac{Post\ baseline\ value}{Baseline\ value}$$

7.1.6 Percentage Change from Baseline

Percentage change from baseline will be defined for post-baseline timepoints as

$$Percentage\ Change\ from\ baseline = \frac{(Post - baseline\ value - Baseline\ value)\ \times 100}{Baseline\ value}$$

If either baseline or postbaseline is missing then percentage change from baseline will also be missing.

7.1.7 Last Contact Date

The last contact date is defined as the maximum date among the following:

- Study Completion Date
- Study Withdrawal Date
- Date of last follow-up visit (only if patient status is not “dead”)
- Date recorded for lost to follow-up
- Date of death.

7.1.8 End of Study

The end of the study is defined as the last patient last visit date at 1 year after last IMP infusion or HSCT (as applicable).

7.1.9 Age at Informed Consent

Age at informed conformed consent (years) will be calculated as:

$$Age\ at\ informed\ consent\ (years) = \frac{Date\ of\ informed\ consent - \ Date\ of\ birth}{365.25}$$

If the date of birth is (partially) missing (e.g. in countries where it is prohibited to provide a complete date of birth), the age (years) is entered in the CRF and will be used directly. Age at informed consent will be rounded to the closest whole year for presentation in the summary tables and listings.

7.1.10 Age at Diagnosis

Age at (initial/current/malignancy) diagnosis (years) will be derived as

$$Age\ at\ diagnosis\ (years) = \frac{Date\ of\ diagnosis - \ Date\ of\ birth}{365.25}$$

If the date of birth is (partially) missing (e.g. in countries where it is prohibited to provide a complete date of birth), the following imputations will be made to allow a calculation of age at diagnosis:

- If only the day is missing, the 15th of the month will be used.
- If day and month are missing, the 1st of July will be used.
- If the date is completely missing, no imputation will be made and age at diagnosis will be left missing.

- If after imputation, the age at diagnosis is greater than age at informed consent, and date of diagnosis is prior to date of informed consent, then the age at informed consent will be used.

Age at diagnosis will be rounded to the closest whole year for presentation in the summary tables and listings.

7.1.11 Time and Temperature Conversions

The following conversion factors will be used to convert days to months or years, or Fahrenheit to Celsius, where applicable:

- 1 month = 30.4375 days
- 1 year = 365.25 days
- 1 week = 7 days
- °C = (°F - 32)/1.8

7.1.12 Medication

7.1.12.1 Medication Related to HSCT

HSCT related medication is defined for Patients with a HSCT date available as follows:

- any medication with “Other” reason for use text containing “HSCT”, “transplant” or “stem cell” (the review of the terms will be performed by the clinical scientist and extended during the study, if applicable) that start no earlier than 16 days before HSCT and end no later than the date of HSCT, i.e.:

  $$Medication\ Start\ Date \geq HSCT\ date - 16\ days \quad and \quad Medication\ End\ Date \leq HSCT\ date$$

  Thereby, if medication start date or end date is (partially) missing, the imputation rules as provided in Section 8.3.2 will be applied prior to making the assessment.
- any medication with “Other” reason for use text containing “Conditioning”.

7.1.12.2 Medication Related to Underlying Malignancy

Medication related to underlying malignancy is identified as any medication with “Other” reason for use text containing “underlying malignancy” or “lymphoma” or “leukemia”, regardless of capitalization of the words.

7.1.12.3 Medication Related to HLH

Medication related to HLH is identified as any medication with reason for use given as “Underlying HLH”.

7.2 Analysis Visit Windows

Besides the visit windows allowed per protocol, time windows are considered in case of missing data for defining study endpoints. For the assessment of endpoints at pre-defined time-points, time windows are summarized in Table 8-1.

In the case that a parameter is not available on the specified study day, the visit windowing is to be applied. If there is more than one observation during the visit window, a value will be selected as follows: the closest value (within the visit window specified in Table 8-1) to the actual study day in question will be chosen first. In the case that there are 2 equidistant values to the study day of interest, the observation that is before the study day in question will be selected. In the case of 2 or more measurements on the same date, the earliest measurement before the infusion will be selected.

For vital signs data, which is recorded at multiple time-points on the same day, the assessment will be considered as a whole when applying visit windowing rules. i.e. if assessments from a certain date would be selected for a timepoint based on the visit windowing rules, but single values from e.g. the 2h

assessment are missing, then this assessment will still be selected and the missing 2h data will not be imputed with the 2h data from another assessment.

Table 8-1 Time Windows for Endpoint Assessment at Pre-defined Time-Points

Timepoints	Time Window used in Analyses
Baseline	Last value prior to or at SD1 (pre-infusion of IMP)
Timepoints during first two weeks of the treatment period (i.e. SD2 up to SD 16) • SD2 • SD4 • SD7 • SD10 • SD13 • SD16	• SD4 ± 1 day • SD7 ± 1 day • SD10 ± 1 day • SD13 ± 1 day • SD16 ± 1 day
Timepoints during the period of bi-weekly IMP infusion: • SD21 (Week 3) • SD28 (Week 4) • SD35 (Week 5) • …	• SD18-24 • SD25-31 • SD32-38 • …
Week 4 Assessment	SDxx-3 to SDxx+5
EOT	SDxx-3 to SDxx+5 (provided before start of conditioning)
Pre-conditioning	Last value prior the start of first conditioning treatment (not earlier than 3 days)
Follow-up Week 1	D3-3 to D3+3
Follow-up Week 2	D10-3 to D10+3
Follow-up Week 3	D17-3 to D17+3
Follow-up Day+30	D30-3 to D30+3
Follow-up Day+60	D60-14 to D60+14
Follow-up Day+100	D100-14 to D100+14
Follow-up 6-months	D180-30 to D180+30
Follow-up 12-months	D360-30 to D360+30
The above follow-up visits are after HSCT. If HSCT is not performed, the time-points are after the last IMP infusion, rather than after HSCT. For the timepoints during the period of bi-weekly IMP administration, SDxx refers to the study day relative to start of IMP. For the Week 4 Assessment timepoint, SDxx refers to the study day corresponding to the date of last IMP in Week 4 + 3 days. For the EOT timepoint, SDxx refers to the study day corresponding to the date of last IMP infusion + 3 days. For the follow-up visits, Dxx refers to the day relative to the date of HSCT, or date of last IMP infusion if HSCT is not performed. It can therefore be a different study day (SD) for each patient.	

7.3 Methods for Missing Data

Imputation of missing data for defining efficacy endpoints is provided in Section 10.7.2. Imputation of incomplete date of birth for calculation of age at initial diagnosis is described in Section 8.1.9. Unless otherwise stated, the data as recorded (including missing) information will be presented in data listings, rather than the imputed data.

7.3.1 Adverse Events

Partially or completely missing dates and missing relationship to IMP or severity for adverse events (AEs) will be imputed as follows. Detailed definitions for AEs are provided in Section 10.8.1.

7.3.1.1 AE Dates

No imputation of (partially) missing AE dates will be made.

Classification of AEs as treatment-emergent will be performed using a conservative approach as described in Section 10.8.1.

7.3.1.2 Relationship and Severity

The relationship to IMP and severity of AEs are assessed by the Investigator and recorded in the CRF.

If the relationship of an AE is missing, it will be assumed as related to IMP for the purpose of identification of Infusion-related reactions (IRR)s and for presentation in summary tables, unless the AE is not treatment-emergent.

If severity of an AE is missing, it will be considered as severe for presentation in summary tables.

7.3.2 Prior and Concomitant Medication

The below imputations of medication start and end dates will be performed to assess if a medication is prior or concomitant. Imputed dates will also be used, as applicable, to identify medication related to HSCT (see Section 8.1.11.1).

(Partially) missing start dates of medications will be imputed as follows:

- If only the day is missing and
    - month/year of the medication start is equal to month/year of first IMP, then the day will be imputed with the day of first IMP.
    - month/year of the medication start is before or after month/year of first IMP, then the day will be imputed with the first day of the month.
- If day and month are missing and
    - year of the medication start is equal to year of first IMP, then day/month will be imputed with the month/day of first IMP.
    - year of the medication start is before or after year of first IMP, then the day/month will be imputed with the 1st January.
- If the start date of the medication is completely missing, then it will be imputed with the date of first IMP intake.

If after the above imputation, the imputed medication start date is after the available medication end date, then the medication start date will be set to the medication end date. If the medication end date is (partially) missing, it will be imputed based on the rules below, before this comparison is made.

(Partially) missing end dates of medications will be imputed as follows:

- If only day is missing, it will be imputed with the last day of the month
- If day and month are missing, they will be imputed with 31st December
- If the date is completely missing, no imputation will be made; the medication is considered as ongoing.

7.3.3 PK and Biomarker Data Imputation

For PK concentration and PD biomarker data, values below the lower limit of quantification (LLoQ) will be presented as “<LLoQ” in the listings.

8.0 Interim Analyses

The study was prematurely terminated and interim analyses, as defined in the protocol version 3.0 (section 11.3.2), will not be performed.

9.0 Statistical Methods

9.1 General Reporting Methods

All statistical analyses will use SAS® version 9.4 or higher.

Categorical data will be presented using counts and percentages, with the number of patients in each category as the numerator for percentages. Percentages will be rounded to one decimal place except 100% will be displayed without any decimal places and percentages will not be displayed for zero counts. With the exception of AE tables or if not stated otherwise, percentages will be based on the number of patients in the respective analysis set with non-missing values. Number of missing values will also be presented but with no respective percentage.

Continuous data will in general be summarized using the number of observations (n), missing, mean, standard deviation, median, minimum, 1st quartile, 3rd quartile, and maximum. Minimum and maximum will be rounded to the precision of the original value. Mean, median, 1st and 3rd quartile will be rounded to one decimal place greater than the precision of the original value. The standard deviation will be rounded to two decimal places greater than the precision of the original value, up to a maximum of three decimal places.

All summarized data will be listed in addition to the summaries described in the sections that follow. The data listings will display nominal visits where applicable i.e. with the exception of Listing 16.2.6.1.1 which shows the Assessment of Overall Response at Week 4/EOT at derived timepoints, the listings will not show the timepoints derived according to the visit windows. Data recorded as “Other” with additional information specified will be presented as “Other: *specified text*” in the listings.
Etiologies without patients recruited will not be presented in the TFLs.

9.2 Patient Disposition

Screen failure data, including reason for screen failure will be listed for all screen failures. Study completion data, including reason for premature withdrawal from study will be listed for the All Treated Analysis Set. The listing will include study drug completion data, including reason for study drug discontinuation.

9.3 Important Protocol Deviations

A listing of all protocol deviations will be presented by patient. Important deviations (i.e. related to study conduct), major deviations (related to statistical analysis) and deviations due to COVID-19 will be flagged in the listing.

The deviation listing will be based on the All Treated Analysis Set.

9.4 Demographic and Baseline Characteristics

Demographic characteristics will be summarized for the All Treated Analysis Set:

- Age (years) and age categories (>= 18 and 64 years, >=65 and 84 years, >= 85 years)
- Sex (Male/Female/Unknown (as per CRF category)); child-bearing potential will be summarized among female patients along with the reason if the patient does not have a child-bearing potential
- Ethnicity (Hispanic or Latino/not Hispanic or Latino/Not Reported as reported on the CRF/Unknown as reported on the CRF)
- Race (American Indian or Alaska Native/Asian/Black or African American/Native Hawaiian or Other Pacific Islander/White/Other)
- Country of the Investigational Site

- Weight (kg) at Baseline
- Height (cm) at Baseline

Demographic characteristics data will be listed for patients in the All Treated Analysis Set. Separate listings will be provided for the following baseline disease characteristics data:

General baseline disease characteristics:

- Date of malignancy diagnosis
- Age (years) at malignancy diagnosis
- Type of malignancy triggering HLH (Lymphoma, Acute Leukemia, Chronic Leukemia, Other Hematologic Neoplasm, Solid Tumor, Other)
- Eastern Cooperative Oncology Group (ECOG) performance status
- Duration (months) between the current and initial HLH diagnosis
- Genetic mutation status and type of mutation identified if present
- Functional test
- Functional test results

Initial baseline disease characteristics:

- Age at HLH diagnosis
- Whether episode of HLH was treated with medication
- Reported response to HLH treatment at initial diagnosis (Complete Response, Partial Response, No Response/Worsening, Other)
- Number of HLH criteria met at initial diagnosis (<5, 5, 6, 7, 8)
- Whether the patients meet each criterion at initial and current HLH diagnosis (separately):
    - Fever (yes/no)
    - Splenomegaly (yes/no)
    - Cytopenias (yes/no)
    - Hypertriglyceridemia and/or hypofibrinogenemia (yes/no)
    - Hemophagocytosis (yes/no)
    - Low or absent natural killer (NK)-cell activity (yes/no)
    - Ferritin >500 µg/L (yes/no)
    - Soluble CD25 >2400 U/mL (yes/no)

9.5 Medical History

(Past) medical history refers to medical conditions that are resolved at or before the Screening Visit. Concurrent (ongoing) medical conditions are those that are ongoing conditions present at Screening.

Medical history and concurrent medical conditions, not specific to HLH diagnosis, will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 22.1.

A listing of medical history including concurrent medical conditions will be provided. Conditions considered by the investigator as HLH-related clinical signs or symptoms will be flagged in the listing.

9.6 Treatments

9.6.1 HSCT

Data from HSCT performed during the study will be listed based on the All Treated Analysis Set, including date or start of conditioning, date of HSCT, stem cell source, graft type, graft manipulation, whether the patient engrafted, platelet/neutrophil engraftment dates, whether a chimerism test was done, date of test, specimen, test result, donor lymphocyte infusion and date performed.

9.6.2 Extent of Study Drug Exposure

The summary of study drug exposure will be based on the All Treated Analysis Set.

The number of doses of study drug received will be calculated as the sum of all IMP administrations recorded on the CRF where the actual dose infused > 0, regardless of whether the dose was completely administered or was interrupted. Number of doses received will be summarized as a continuous variable using descriptive statistics and by category (<3 doses; ≥3 and <7 doses; ≥7 doses and <11 doses, ≥11 doses).

The infused dose (mg/kg) at each visit will be calculated from total dose infused (mg) at the visit adjusted for the patient's weight as

$$Infused\ dose\ (mg/kg) = \frac{Total\ Dose\ Infused\ (mg)}{Weight\ (kg)}$$

The patient's weight recorded at the timepoint closest, but prior, to the infusion date will be used for the calculation. The maximum dose infused (mg/kg) for each patient is determined as the maximum infused dose (mg/kg) per patient administered throughout the trial. Maximum dose infused will be summarized by category (≤3 mg/kg; >3 mg/kg and ≤6 mg/kg; >6 mg/kg and ≤10 mg/kg; >10 mg/kg).

Total treatment duration (days) will be calculated from the dates of dosing as recorded on the CRF as:

$$Treatment\ duration\ (days) = Date\ of\ last\ dosing - Date\ of\ first\ dosing + 1.$$

Total treatment duration (weeks) will be calculated as:

$$Treatment\ duration\ (weeks) = \frac{Treatment\ duration\ (days)}{7}.$$

Total treatment duration will be summarized by duration category in weeks (<4 weeks; 4 to <8 weeks; 8 to <12 weeks; ≥12 weeks). Treatment duration (days) will also be summarized as a continuous variable using descriptive statistics.

Additionally, the following summaries of treatment exposure will be provided:

- Maximum dose prescribed (≤3 mg/kg; >3mg/kg and ≤6mg/kg; >6mg/kg and ≤10mg/kg; >10mg/kg)
- Minimum dose prescribed (≤3mg/kg; >3mg/kg and ≤6mg/kg; >6mg/kg and ≤10mg/kg; >10mg/kg)
- Cumulative administered dose in mg/kg as a continuous variable

Study drug exposure data will be listed for patients in the All Treated Analysis Set.

9.6.3 Prior and Concomitant Medications

Medications that started prior to the first IMP administration will be considered prior medications. Any medications, continuing or starting after the date of the first IMP administration will be considered to be concomitant. If a medication starts prior to the date of first dose and continues after the date of first dose it will be considered both prior and concomitant. In the event of missing or partially missing medication start date, imputation rules as provided in Section 8.3.2 will be used prior to making the assessment.

Prior and concomitant medications will be coded using version Sep 2019 of the World Health Organization (WHO) Drug Dictionary.

All prior and concomitant medication data will be presented in data listings, with ATC levels 2 and 4 and PT, based on the All Treated Analysis Set, including a reason for use category.

Blood transfusion product data will also only be listed.

9.7 Efficacy Analyses

Any Efficacy data listed will be presented in data listings using the All Treated Analysis Set.

9.7.1 Hypothesis Testing Strategy and Multiplicity

No statistical tests will be performed, adjustment for multiplicity is not applicable.

9.7.2 Imputation Methods for Overall Response

The values for some parameters defining the primary efficacy endpoint (Overall Response) may be missing, thus impacting the response assessment. Imputation rules for assessing Overall Response at Week 4 Assessment Visit, EOT, and throughout the study have been pre-defined as follows:

Response at Week 4 Assessment Timepoint

- Missing data at Week 4 Assessment (i.e. 3 days after the last Week 4 infusion of IMP) for the parameters relevant to the assessment of Overall Response will be imputed by the closest available value within -3 and +5 days before and after the last IMP date in Week 4 + 3 days, respectively, and before HSCT date. For body temperature, in the event of repeated measures recorded on the same study day, the highest value will be selected.
- No imputation will be applied if no value is available within the analysis time window. That parameter will be viewed as missing and complete response will not be assigned even if all other parameters are normal.
- Values recorded at both scheduled and unscheduled visits will be considered.

Response at EOT

- Missing data at EOT (i.e. 3 days after the last infusion of IMP) for the parameters relevant to assessment of Overall Response will be imputed by the closest available value within -3 and +5 days before and after the EOT date respectively (provided upper window limit is before the date of HSCT). For body temperature, in the event of repeated measures recorded on the same study day, the highest value will be selected.
- No imputation will be applied if no value is available within the -3/+5 time window. That parameter will be viewed as missing and complete response will not be assigned, even if all other parameters are normal.
- Values recorded at both scheduled and unscheduled visits will be considered.

Response Throughout the Study for Time to First Response

- For the secondary endpoint time to first response), both scheduled and unscheduled visits will be used for assessment of response.
- If a parameter needed for the response assessment is missing, the midpoint approach will be used to estimate the missing value, i.e., the last observed value before the windowed timepoint with a missing value will be averaged with the first value available after the missing period. Imputed values will be flagged in the listing. This rule will be applied to parameter values missing for a period of up to a maximum of 7 consecutive days i.e. the date of the first non-missing value available after the windowed timepoint with the missing value – the date of the last observed value before the windowed timepoint with a missing value +1 must be <= 7 days for the average to be used. In case that, after imputation, data is still missing in any of the criteria considered for the definition of response, a Complete Response cannot be assigned.

Since data relevant to CNS assessment has been collected at irregular time-points (based on individual patient conditions), the assessment of CNS status will be carried-forward until the time-points of

relevance for response analysis (e.g. a CNS improvement assessed at SD7 will be considered for ORR at Week 2 if CNS assessment was not performed at Week 2).

Data from different dates may contribute to the assessment of response. Therefore, the date of response or non-response is derived as the maximum of all dates within the above specified imputation windows.

9.7.3 Primary Endpoint Evaluation

9.7.3.1 Primary Analysis

The primary endpoint will be based on the Week 4 Assessment Visit or EOT (whichever is earlier) with missing data imputed as described in Section 10.7.2. In case of missing data for Week 4 Assessment Visit that was performed (e.g. because one parameter is missing even after imputation), the EOT timepoint and data will not be used for analysis. Patients without a Week 4 Assessment Visit or EOT value (after imputation) are considered as non-responders.

No analysis will be performed on the primary endpoint; it will only be included in a listing of overall response at Week 4/EOT for the All Treated Set. A second listing for overall response at the scheduled visits (as opposed to the listing showing derived timepoints above) will be presented at Week 4/EOT for the All Treated Set.

9.7.4 Secondary Analyses

9.7.4.1 Best Response on Treatment

The best response on treatment, i.e., Complete Response, Partial Response, or No Response will not be presented.

9.7.4.2 Overall Response at EOT

The Overall Response at EOT will not be analyzed. Missing data will be imputed as described in Section 10.7.2; it will only be included in a listing of overall response at Week 4/EOT for the All Treated Set.

9.7.4.3 Overall Survival

Overall Survival is defined as the time from the date of first dose of IMP to the date of death from any cause and will be listed using time-to-event methodology. For this analysis, patients alive at the end of the study or lost-to-follow-up will be censored at the last known date alive. Alive patients who prematurely withdraw prior to the end of study from the study will be censored at the withdrawal date. Missing data will be imputed as described in Section 10.7.2. Overall survival will be included in a listing of time-to-event efficacy endpoints for the All Treated Analysis Set. Data from the survival status CRF page will be listed for the All Treated Analysis Set.

9.7.4.4 Time to Complete or Partial Response

The time to first response (CR or PR) is defined as the time from the date of first dose of IMP to the first achievement of response. Patients without a response at the end of the study or lost-to-follow-up will be censored at the last known date the patient is known to not have had a response, using the rules defined in Section 10.7.2. Time to first response will be included in a listing of time-to-event efficacy endpoints for the All Treated Analysis Set.

9.7.4.5 Duration of Response

The duration of first response will not be presented.

9.7.4.6 HLH Relapse

HLH Relapse will not be presented.

9.7.5 Other Efficacy Variables

Investigator assessment of clinical response will be listed for the All Treated Analysis Set.

9.8 Safety Analyses

All safety analyses including AEs will be performed using the All Treated Analysis Set.

In all summaries of AEs, counting will be by patient, not event, and patients are only counted once within each relevant category. If not stated otherwise, summaries of AEs will also include the number of events reported and all events will then be counted.

Summaries of AEs by SOC and PT will, in general, be sorted in descending frequency of SOC and of PT within SOC. If the number of patients is the same within one PT or SOC, then these PTs or SOCs will be sorted by descending number of events (if applicable), then alphabetically, (PTs within the SOC).

If not stated otherwise, only Treatment-Emergent Adverse Events (TEAEs) will be included in the summaries; listings will include all relevant AEs and TEAEs will be flagged.

9.8.1 **Adverse Events**

An adverse event is any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal (investigational) product, whether or not related to the medicinal (investigational) product (see the ICHE2A Guideline for Clinical Safety Data Management: Definitions and Standards for Expedited Reporting). Adverse Events will be coded using MedDRA version 22.1.

Any AE that occurs after the start of first IMP administration, or any pre-existing condition which increases in severity after the start of first IMP administration, will be considered as a TEAE. In the case of partially or completely missing AE start date/time, the AE will be considered treatment emergent, unless the available date/time information clearly indicates that the AE started prior to start of first IMP administration as follows:

- If the start time is missing, but start date is before start date of IMP administration.
- If the day of the AE is missing, but the year/month are before the year/month of the start of IMP administration.
- If the day and month of the AE are missing, but the year is before the year of the start of IMP administration.
- If the stop date of the AE is before the start date of the first IMP administration.

Infections are defined as all PTs included in the SOC “Infections and infestations” and/or in the HLGT “Microbiology and serology investigations”.

IRRs are defined as any AE that is reported to have occurred within 24 hours after start of any study drug infusion and assessed as related to study treatment by the Investigator, but excluding AEs coded to any of the following SOCs:

- Infections and infestations
- Congenital familial and genetic disorders
- Neoplasms benign, malignant and unspecified (incl. cysts and polyps)
- Product issues
- Social circumstances
- Surgical and medical procedures

If the onset time of the AE or the start time of the infusion is missing, then an AE with an onset date equal to an infusion date or infusion date + 1 will be considered for the assessment of IRRs, unless the available information clearly indicates that the AE was not treatment-emergent.

The imputation methods defined in Section 8.3.1 will be used for classifying severity of AE or relationship to IMP as appropriate.

A summary for all AEs will be provided that includes number and percentage of patients and number of events for the following:

- All AEs (TEAEs, pre-treatment AEs)
- TEAEs related to study drug
- Severity of TEAEs (Mild, Moderate, Severe)
- TEAEs leading to discontinuation of study drug
- TEAEs resulting in death
- Serious TEAEs
- Non-serious TEAEs
- Serious TEAEs related to study drug
- IRRs
- Serious IRRs
- Treatment-emergent infections
- Severity of treatment-emergent infections (Mild, Moderate, Severe)

Summaries of the number of events, and number and percentage of patients reporting:

- Any TEAE
- Any non-serious TEAE
- Any treatment-related TEAE
- Any TEAE leading to study drug discontinuation
- Any IRR
- Any treatment-emergent infection

categorized by SOC and PT will be presented.

All AEs (including non-treatment-emergent events) recorded on the CRF will be listed. Separate listings will be provided for TEAEs leading to discontinuation of study drug, IRRs, and Infections.

9.8.2 Deaths and Serious Adverse Events

Summaries of the number of events, and number and percentage of patients reporting:

- Any serious TEAE
- Any serious TEAE leading to study drug discontinuation
- Any treatment-related serious TEAE
- Any serious TEAE resulting in death

categorized by SOC and PT will be presented.

All SAEs and AE leading to death will be presented in the listings.

9.8.3 Infections

In case of suspicion of infection, the following searches for infections will be performed:

- Adenovirus
- Epstein-Barr virus (EBV)
- Cytomegalovirus (CMV)

- Atypical mycobacteria
- Histoplasma Capsulatum
- Leishmania

Tests for Tuberculosis are further performed every 4 weeks throughout the study. Tests for human immunodeficiency virus (HIV), Hepatitis B and Hepatitis C are performed at Screening. Results from each test for infection performed are recorded on the CRF and all Infection test data will be listed based on the All Treated Analysis Set.

9.8.4 Laboratory Data

Routine laboratory assessments include Hematology (Complete Blood Cell Count), Biochemistry, Urinalysis, and Coagulation and will be performed at screening, Visit 1 (pre-infusion) and at each post-baseline visit, including visits during the follow-up period.

Laboratory tests will be performed locally, and results will be recorded on the CRF.

Laboratory values will be classified as within range/normal, abnormal high, or abnormal low, based on the local laboratory's normal ranges for the respective parameter.

All laboratory data, including coagulation parameters and pregnancy test results will be listed. Listings will include data recorded at scheduled and unscheduled assessments. The listings will present the results in standard units and the ranges will be converted to standard units for the purpose of presentation in the listings. Laboratory values outside the reference ranges will be flagged in the listings. If there is no matching range to the original unit is available in the raw data then the range will be missing in the listing. There will be separate listings for each of the following laboratory test groups:

- Hematology Parameters: Red Blood Cells, Hemoglobin, Hematocrit, White Blood Cells (WBC), Neutrophils (% and absolute count), Lymphocytes (% and absolute count), Platelets
- Clinical Biochemistry Parameters: Albumin, Alkaline Phosphatase, Alanine Aminotransferase (ALT)/serum glutamic-pyruvic transaminase (SGPT), Aspartate aminotransferase (AST)/serum glutamic-oxaloacetic transaminase (SGOT), Blood Urea Nitrogen (BUN), Glucose, Lactate Dehydrogenase, Serum Creatinine, Total Bilirubin, Total Protein, Triglycerides, C-Reactive Protein (CRP), Ferritin
- Urinalysis Parameters: Urine Glucose, Urine Protein, Urine Ketones, Urine Blood, Urine Leukocytes
- Coagulation Parameters: Activated Partial Thromboplastin Time (APTT), Prothrombin Time, Prothrombin Time INR (PTINR), D-Dimers, Fibrinogen

For fibrinogen and D-Dimer, the following categories will be flagged in the listing in addition to any local ranges:

- Fibrinogen: Abnormal (Low): ≤150 mg/dl; Normal (Within): > 150 mg/dl
- D-dimer: Abnormal (High): ≥ 1 µg/ml; Normal (Within): < 1µg/ml

9.8.5 Vital Signs

Vital signs will be recorded at each visit and measured at pre-infusion, at the end of infusion, and 1 and 2 hours after the end of infusion, except for body temperature which will only be measured pre-infusion.

Vital signs recorded at scheduled and unscheduled assessments (including body temperature (Celsius), systolic and diastolic blood pressure (mmHg), heart rate (beats/min), and respiratory rate (breaths/min)) will be presented in the listings. The results for temperature will be presented as reported and the correction depending on the location will not appear in the listing.

9.8.6 Physical Examinations, ECGs, and Other Observations Related to Safety

Physical examination parameters include General appearance, Thorax/Lungs examination, Cardiovascular examination, Skin examination, Musculoskeletal examination, Neurological examination, and Abdominal examination (including liver size and spleen size). All physical examination results will be presented in a data listing.

Scheduled ECG assessments are planned during Screening and at EOT. All recorded ECG data will be presented in a data listing.

All recorded imaging data, including abnormal ultrasounds, chest x-rays, and brain MRIs will be presented in data listings.

Hospitalization data will be listed.

9.9 PK/PD Analyses

PK/PD modelling analyses will not be performed. For the purpose of the present SAP, the PK concentration data and PD biomarkers data (CXCL9, sCD25, total IFNγ, IL-6) will be listed for the All Treated Analysis Set. ADA data will be listed for the All Treated Analysis Set.

10.0 Glossary of Abbreviations

Glossary of Abbreviations:	
ADA	Anti-drug antibodies
AE	Adverse event
ALT	Alanine Aminotransferase
APTT	Activated Partial Thromboplastin Time
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutical Class
BUN	Blood Urea Nitrogen
CDISC	Clinical Data Interchange Standards Consortium
CNS	Central nervous system
CR	Complete Response
CRF	Case Report Form
CRP	C-Reactive Protein
CSF	Cerebrospinal fluid
CSR	Clinical Study Report
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EOT	End of Treatment
G-CSF	Granulocyte-colony-stimulating factor
HIV	Human immunodeficiency virus
HLGT	High Level Group Term
HLH	Hemophagocytic lymphohistiocytosis
HSCT	Hematopoietic stem cell transplantation

iDMC	Independent Data Monitoring Committee
IL-6	Interleukin 6
IMP	Investigational medicinal product
IRR	Infusion-related reaction
LLoQ	Lower limit of quantification
KM	Kaplan-Meier
M-HLH	Malignancy-associated hemophagocytic lymphohistiocytosis
ORR	Overall Response Rate
PD	Pharmacodynamic
PK	Pharmacokinetic
PR	Partial Response
PT	Preferred Term
PTINR	Prothrombin Time INR
RBC	Red Blood Cells
SAP	Statistical Analysis Plan
SAE	Serious adverse event
sCD25	Soluble CD25 (i.e., soluble IL-2 receptor)
SD	Study Day
SGPT	Serum glutamic-pyruvic transaminase
SGOT	Serum glutamic-oxaloacetic transaminase
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
WBC	White Blood Cells

11.0 Appendices

11.1 Overall Response Detailed Derivation

The HLH parameters used to assess disease response according to the definition provided in Table 7-1 are described in Table 12-1. Details for deriving overall response are described in Table 12-2

Table 12-1 HLH Parameters for Assessment

HLH Parameter	Definition of Normalization
Parameters that are measured for improvement	
Body temperature	Below 37.5°C in all measurements performed on that day
Spleen as assessed at physical examination	Reported as Normal (= 0 cm from costal margin)
Absolute Neutrophil Count and G-CSF administration	Absolute Neutrophil Count is equal/greater than 1.0 x 10^9/L, and no G-CSF (ATC code=L03AA) has been administered in the previous 4 days, except for medication given with reason for use as “Medical History” and associated Medical History Term is not HLH-related.

Platelet count and platelet transfusion	Platelet count is equal/greater than 100 x 10^9/L, and no platelet transfusion administered in the previous 4 days
Ferritin	Less than 2000 µg/L
Fibrinogen and D-dimers	Fibrinogen greater than 1.5 g/L and/or D-dimers equal/less than 500 ug/L
CNS disease involvement, as assessed by medical team	Normal results or improvement in all three of the following • Physical examination • CSF • Brain MRI Normal ranges for CSF are as follows: • White Blood Cells < 5 per mm^3 • Red Blood Cells = 0 • Protein < 0.5g/l • Neopterin 9–20 nmol/l
Parameters that are measured for worsening of disease	
sCD25 levels	Fold Change from Baseline (for the last two sCD25 data points before assessment) is not > 2

Table 12-2 Derivation of Overall Response

Response	Application of Criteria	Example
Complete Response	All HLH parameters that were abnormal at baseline must be normalized. All other parameters that were normal at baseline must still meet the definition of normalized.	Subject started with 5 abnormal parameters. All 5 improved, and the remaining parameters continue to meet definition of normalized.
Partial Response	At least 3 of the HLH parameters that were abnormal at baseline must be normalized or improved (i.e., >50% change from baseline). The remaining may be the same or worse as baseline, unless 2 or more parameters that were normal at baseline no longer meet the definition of normality. Note: If 3 parameters were abnormal at baseline, only 2 need to be at least improved. If 2 parameters were abnormal at baseline, only 1 needs to be improved. Definition of 50% improvement from baseline: • Spleen size decreased by 50%, as recorded in cm from costal margin at physical examination • Absolute Neutrophil Count increased by 50%, and G-CSF has not been administered in the previous 4 days • Platelet count increased by 50%, and no platelet transfusion has been administered in the previous 4 days • Ferritin decreased by 50% • Fibrinogen increased by 50% or D-Dimer decreased by 50% • CNS= 'Improved'.	Subject started with 4 abnormal parameters. Two (2) of the parameters now meet definition of normalized. The third improved >50%, but is still abnormal. The fourth is still abnormal and did not improve >50%. No more than 1 parameter (normal at baseline) crosses the normality threshold.

No Response	If a subject does not meet the criteria for at least Partial Response, response is categorized as “No Response”	Patient who are not meeting any PR or CR criteria.

11.2 Planned Table, Listings, and Figures

A Table of Content with all TLFs planned for this study is provided in a separate document.