STATISTICAL ANALYSIS PLAN

AB2 Bio Ltd

NLRC4/XIAP.2016.001

Protocol Title:	Multicenter, double-blind, placebo-controlled, randomized withdrawal trial with Tadekinig alfa (r-hIL-18BP) in patients with IL-18 driven monogenic auto-inflammatory conditions: NLRC4 mutation and XIAP deficiency
Protocol Version and Date:	Version 6.0; 3 September 2021
Sponsor:	AB2 Bio Ltd EPFL Innovation Park Building B, 4th Floor CH-1015 Lausanne, Switzerland
Prepared By:	[redacted] Precision for Medicine 6005 Hidden Valley Road, Suite 170 Carlsbad, CA 92011
Document Version and Date:	Version 7.0; 25 September 2023

1 STATISTICAL ANALYSIS PLAN APPROVAL

Sponsor:	AB2 Bio Ltd
Clinical Protocol Number:	NLRC4/XIAP.2016.001
Protocol Title:	Multicenter, randomized, double-blind, placebo-controlled trial with Tadekinig alfa (r-hIL-18BP) in patients with IL-18 driven monogenic auto-inflammatory conditions: NLRC4 mutation and XIAP deficiency
Document File Name:	AB2Bio_NLRC4-XIAP.2016.001_v7.0_25Sep2023.pdf
Document Version and Effective Date:	Version 7.0; 25 September 2023

Approved By:

25-Sep-2023

[redacted]
Medical Director
AB2 Bio Ltd.

Date

26-Sep-2023

[redacted]
Precision for Medicine

Date

2 TABLE OF CONTENTS

3 LIST OF ABBREVIATIONS

Table 1 List of Abbreviations

Abbreviation	Definition
ADA	anti-drug antibody
AE	adverse event
AESI	adverse events of special interest
ATC	Anatomical Therapeutic Chemical
BMI	body mass index
CARD	Caspase Recruiting Domain
CMV	cytomegalovirus
CNS	Central Nervous System
CRP	C-reactive protein
CSR	clinical study report
DSMB	Data Safety Monitoring Board
EBV	Epstein-Barr virus
eCRF	electronic Case Report Form
EOS	End of Study
ESR	erythrocyte sedimentation rate
ET	Early Termination
FAS	Full Analysis Set
GC	glucocorticoids
GI	gastrointenstinal
ICH	International Council for Harmonisation
IL-18	interleukin-18
IL-18BP	Interleukin-18 binding protein
IMP	Investigational Medicinal Product
IRR	incident rate ratio
KM	Kaplan-Meier
LOCF	last observation carried forward
mAIDAI	modified Auto-Inflammatory Disease Activity Index
MAS	Macrophage Activation Syndrome
MedDRA	Medical Dictionary for Regulatory Activities
NLRC4	Nucleotide-binding oligomerization domain, leucine rich repeat and CARD domain containing 4
PEY	Person Exposure Years
PGA	Physician Global Assessment
PK	Pharmacokinetics
PKAS	PK Analysis Set

Abbreviation	Definition
PPS	Per-Protocol Set
Q1	25th percentile
Q3	75th percentile
r-hIL-18 BP	Human Recombinant Interleukin-18 Binding Protein
RDBPC	Randomized Double-Blind Placebo-Controlled
RW	Randomized Withdrawal
RWS	Randomized Withdrawal Analysis Set
SAE	serious adverse event
SAOL	Single Arm Open-Label
SAP	statistical analysis plan
SAS	Safety Analysis Set
s.c.	subcutaneously
SI	Système International
SOC	standard of care
TA	Tadekinig alfa
TEAE	treatment-emergent adverse event
ULN	upper limit of normal
V	Visit
WHO	World Health Organization
XIAP	X-linked inhibitor of apoptosis

4 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide comprehensive and detailed descriptions of the methods and presentation of data analyses proposed for AB2 Bio Ltd Protocol NLRC4/XIAP.2016.001 (Multicenter, placebo-controlled, randomized withdrawal trial with Tadekinig alfa [r-hIL-18BP] in patients with interleukin-18 [IL-18] driven monogenic auto-inflammatory conditions: NLRC4 mutation and XIAP deficiency). Descriptions of planned analyses are provided in order to avoid post hoc decisions that may affect the interpretation of the statistical analysis. The statistical methods applied in the design and planned analyses of this study are consistent with the International Council for Harmonisation (ICH) guideline *Statistical Principals for Clinical Trials* (E9) (1998).

This SAP will be finalized prior to data analysis, before breaking the treatment blind on the Randomized Withdrawal (RW) phase, and prior to database lock to provide full details, including templates for tables, listings, and figures, to be presented in the clinical study report (CSR). Any changes between the statistical methods provided in the clinical study protocol and this SAP will be explained herein; any changes or deviations from this SAP relative to the final analysis will be fully documented in the CSR. Minor changes or deviations from the templates for tables, figures, and listings need not be documented in the CSR.

5 STUDY OBJECTIVES

5.1 Primary Study Objectives

The primary objectives of this study are:

- To assess clinical efficacy of Tadekinig alfa (TA) in monogenic auto-inflammatory diseases with ongoing inflammation and deleterious mutations of Nucleotide-binding oligomerization domain, leucine-rich repeat and caspase recruiting domain (CARD domain) containing 4 (NLRC4) – Macrophage activation syndrome (MAS) mutation or X-linked inhibitor of apoptosis (XIAP) deficiency
- To assess laboratory/biological evidence of efficacy

5.2 Secondary Study Objectives

5.2.1 Secondary Efficacy Objective

The secondary efficacy objective of the study is to assess the main clinical features and laboratory biomarkers (ie, characteristics of each patient profile) longitudinally during the single arm open-label (SAOL) phase for each experimental patient.

5.2.2 Secondary Safety Objective

The secondary safety objective of the study is to assess the safety and tolerability of TA treatment in monogenic, autoinflammatory disease harboring deleterious mutations of NLRC4-MAS or XIAP.

6 INVESTIGATIONAL PLAN

6.1 Overall Study Design

This is a Phase 3 study to assess the safety and efficacy of TA in patients with monogenic, IL-18 driven auto-inflammation due to NLRC4-MAS mutation or XIAP deficiency. Because of the likelihood for pathogenic IL-18 in certain monogenic diseases, patients known to harbor deleterious mutations in NLRC4-MAS or XIAP and who have a history of ongoing inflammation will be enrolled if they have ferritin $\geq$ 500 ng/mL or persistent C-reactive protein (CRP) elevation $\geq$ 2 times the upper limit of normal (ULN) and the patients have a Modified Auto-Inflammatory Disease Activity Index (mAIDAI) $\geq$ 4.

This study is designed with SAOL of TA treatment duration for 18 weeks followed by a 16-week RW phase for efficacy and safety evaluation, with no interruption between the two phases of treatment. All patients who complete the 18 weeks of treatment during the SAOL phase and showing a partial or complete response at Week 18 (V8) will be randomized to one of the two treatment arms in a 1:1 ratio to begin treatment in the RW phase. Patients will receive the study treatment (placebo or TA) for up to an additional 16 weeks. During the RW phase, patients who develop a partial or full disease reactivation before the 16-week treatment is completed will proceed to the End of Study visit and complete the RW phase. The screening period will occur before the SAOL phase and before the first dose of Investigational Medicinal Product (IMP).

The physicians and the patients (or legal representative) will have the option to discontinue treatment if they feel that pursuing the treatment is not appropriate. The physician can decide any time to start rescue immunosuppression and therefore discontinue the patient from the study. In this case, all key safety and efficacy outcomes will be followed up. Patients will continue to come to the scheduled visits for safety and efficacy assessments. The occurrence of adverse events of severity grades 3 or 4 may result in transient or permanent treatment discontinuation. If treatment interruption is more than 3 weeks, the patient will be permanently discontinued from the study.

An estimated maximum of 15 patients will be enrolled in this study at approximately 11 clinical sites in the United States, Canada, and Germany. A maximum of 15 patients enrolled will ensure the randomization of at least 10 patients in the RW phase.

6.1.1 Screening Period

The screening visit (V0) will occur before the start of the study treatment. At V0, the patient and/or the patient's legal guardian(s) will sign a Child Assent Form and/or an

Informed Consent Form that will be collected by the Investigator. All eligibility criteria will be reviewed during the screening period AND must all be met in order to determine patient eligibility into the study.

6.1.2 Single Arm Open-Label Phase

Eligibility criteria should be confirmed again at the Baseline Visit (V1). Patients experiencing an active disease will begin TA treatment at V1.

Beginning at V1, TA will be administered in addition to the standard of care (SOC) treatment used for the control of the active disease. In the absence of an approved medication for this disease, the SOC as defined in the study protocol consists of glucocorticoids (GC), with a 28-day tapering schedule as follows:

- 28-days (4 weeks): Daily PO/IV Dexamethasone 10 mg/m^2 x 7 days, 5mg/m^2 x 7 days, 2.5 mg/m^2 x 7 days, 1.25 mg/m^2 x 7 days.

Patients who respond to the initial 28-day course of TA + GC (partial or complete response as defined in Section 6.4.1.2) will receive TA for 14 additional weeks.

Patients who respond to TA + SOC initial course and have a disease reactivation again during the SAOL phase will receive another course of TA + SOC with the same GC dosage and tapering schedule used for the initial treatment of the active disease. Maintenance treatment, with the permitted medications, at stable and at the same doses utilized at study entry are allowed. Dose increments for the permitted co-medications or their introduction after the beginning of the study are not permitted and represent treatment failures and study early termination. If permitted co-medications are weaned off during the course of the study and the investigator decides to re-introduce them, they are permitted as long as the doses are the same or lower than the doses used at study entry.

Upon lack of response to combined treatment after enrollment or upon the occurrence of a disease reactivation during the course of the SAOL after initial response, the investigator may decide to treat the patient with rescue immunosuppressors. This decision results in patient discontinuation from the study.

Tadekinig alfa will be administered at 2 mg/kg (with a maximum of 160 mg) subcutaneously (s.c.) for up to approximately 18 weeks.

At V0 and V1, patients will undergo a physical examination, measurement of vital signs, and routine laboratory testing (including hematology, clinical chemistry, urinalysis, erythrocyte sedimentation rate [ESR], CRP, and coagulation panel). Biological evaluation and genetic diagnosis of auto-inflammatory conditions (NLRC4-MAS, XIAP deficiency, and other monogenic associations with MAS) will be done at V0 (screening visit). Samples for flow cytometry for suspected XIAP deficiency will be analyzed during the screening period and prior to V1. Complete medical and disease history will also be collected during V0. In addition, prior to initiation of therapy, cytomegalovirus (CMV)

and Epstein-Barr virus (EBV) tests, serology tests, and a complete infectious workup will be performed.

Specific laboratory determinations such as total and free IL-18 and IL-18BP will be performed at V1 and subsequent visits.

The global disease activity will be assessed by the mAIDAI. The mAIDAI will be applied at all study visits and will be available for review throughout the study. For the evaluation of response to therapy and disease reactivation, only the objective parameters of the mAIDAI will be taken into consideration. A Physician Global Assessment (PGA) tool will also be applied at all study visits by the physician to assess the global disease status. The patient's or caregiver's evaluation of the health status questionnaire will be applied at V1 and the end of the SAOL phase.

Assessments during the SAOL phase (V1 to V8) will include complete physical examination, review of all adverse events (AEs), and routine and specific laboratory analysis. Samples for anti-drug antibodies (ADAs) will be collected at Baseline of SAOL (V1), Week 2 (V3), Week 4 (V5), Week 12 (V7), and Week 18 (V8).

6.1.3 Randomized Withdrawal Phase:

All patients who have completed 18 weeks of treatment during the SAOL phase and showing a partial or complete response at Week 18 (V8) will be randomized in a 1:1 ratio to begin treatment in the RW phase. Patients will receive the study treatment (placebo or TA) for up to an additional 16 weeks. Experiencing a disease reactivation at any point during the 18-week SAOL phase will not prevent a patient from enrollment into the RW phase; rather, patients are required to meet the criteria for ongoing TA treatment throughout the SAOL phase and show a partial or complete response at Week 18 (V8).

Treatment interruption is not recommended between the two phases. Baseline data for the RW phase will be the last data collected prior to randomization at the Week 18 (V8) visit.

During the RW phase, visits will be scheduled at Week 20 (V9), Week 22 (V10), and Week 26 (V11), Week 30 (V12), and Week 34 (V13/End of Study [EOS]). At RW visits, patients will undergo a complete physical examination and laboratory testing (including hematology, clinical chemistry, urinalysis, ESR, CRP, ferritin, and coagulation panel), and determinations of total and free IL-18 and of IL-18 BP. The mAIDAI and PGA will be collected at all visits in the RW phase. The patient's or caregiver's evaluation of the health status questionnaire will be applied at the end of the RW phase. Samples for ADAs will be collected at Week 22 (V10), Week 26 (V11), and Week 34 (V13/EOS).

During the RW phase, patients who develop a partial or full disease reactivation before the 16-week RW treatment period is completed will proceed to the End of Study visit and complete the RW phase. The investigator may also decide to start rescue immunosuppression and discontinue treatment. If applicable, data of administration of other medications or rescue immunosuppression during both phases (SAOL and RW) will be collected, including the time of treatment start and finish and whether or not there was a response to the rescue treatment.

Patients who discontinue treatment prior to the end of either treatment phase (SAOL or RW) will be followed for all key efficacy and safety assessments up to the End of Study visit.

6.1.4 Additional Standard of Care Cycles or Rescue Immunosuppression

A) If no response after the first combined treatment course (SOC + IMP) at the initiation of the SAOL phase:

If patients do not respond to study treatment after the initial combined treatment course, they will be treated with rescue immunosuppression or with new/ increased doses of permitted concomitant treatments. These cases will be considered treatment failures during the SAOL phase and will exit the study.

B) Patients with a disease reactivation after having responded to therapy in the SAOL phase:

If patients show signs of systemic inflammatory reactivation (ie, partial or full disease reactivation) after having responded to therapy in the SAOL phase phase, they will receive another cycle of SOC in addition to the IMP.

C) Patients with a disease reactivation during the tapering down of SOC during the combined treatment in the SAOL phase:

If after an initial improvement in response parameters, the patients show signs of increasing inflammation (ie, partial or full disease reactivation) during the tapering down of SOC while in combined treatment, another full cycle of extended SOC (4 weeks) in addition to the IMP will be initiated at the original dose.

In the last cases (B or C), the Investigator may also decide to start rescue immunosuppression, and in this case, the patient should discontinue study treatment due to treatment failure.

Any patient starting rescue immunosuppression treatment will be considered a treatment failure and will proceed to the study discontinuation/early terminations procedures. Data related to response to the rescue treatment will be collected including all key efficacy and safety endpoints.

It is recommended that patients suffering from severe CNS manifestations are put on rescue immunosuppression treatment.

Patients that present a partial or full disease reactivation after previous response to SOC + TA in the SAOL can remain in the study and receive another cycle of SOC in addition to the continuous TA treatment if, according to the treating physician, the severity does not require rescue immunosuppression. Patients that present a partial or full disease reactivation in the RW Phase before the 16-week treatment is completed will proceed to the End of Study Visit and complete the RW Phase.

6.1.5 Original Study Design

The study under the original protocol (20 September 2016) and through Amendment 2 (Version 3.0, 07 April 2017) was designed with an 18-week Randomized Double-Blind Placebo-Controlled (RDBPC) phase, followed by an 8-week RW phase for further efficacy and safety evaluation. Three patients were enrolled in the study under these protocol versions and each either completed the RDBPC phase and RW phase or terminated the study early.

During the RDBPC phase of the study, patients were randomized to TA or placebo at the Baseline/Randomization Visit (V1) in a 1:1 ratio. Patients randomized to TA during this phase effectively enrolled in the same study design as described for Protocol Amendment 4 and following Amendments. See Section 7.1.6 for a description of how these patients will be considered for analysis purposes. The treatment blind for the RW phase for patients enrolled under the original study design will be maintained until the blind is broken for the RW phase of the updated study design, once all patients have completed or terminated the study and the database is locked.

Versions 5.0 and later of the protocol extended the RW phase to 16 weeks and specified that patients who presented a partial or full disease reactiviation in the RW Phase before the 16-week treatment is completed would come off study and proceed to the End of Study Visit.

6.2 Schedule of Assessments

For the complete schedule of assessments, refer to Section 7 (Schedule of Events) of the clinical study protocol.

6.3 Treatments

6.3.1 Treatments Administered

Tadekinig alfa is a soluble glycoprotein of [redacted] amino acids produced from Chinese Hamster Ovary cell line. The polypeptide chain contains [redacted] The average molecular weight of the full-length polypeptide moiety of recombinant human IL-18BP (r-hIL-18BP), calculated on the basis of the amino acid composition, is approximately [redacted] kD. The relative molecular mass of the whole molecule is approximately [redacted] kD (including glycans).

Tadekinig alfa is supplied as a colorless to slightly yellow, sterile solution for injection in glass vials containing [redacted] as excipients. It is available in a concentration of 20mg/0.5mL.

A sodium chloride [redacted] will be used as placebo. To ensure that the IMP remains blinded during the RW phase, both the actual

drug and the placebo solutions will be supplied in identical vials and similar labeling and will be indistinguishable in terms of their texture, color, and smell.

Patients will receive a dose of 2 mg/kg (with a maximum of 160 mg) of TA s.c. every 2 days ± 5 hours for a total treatment period of 18 weeks in the SAOL phase.

Patients in the RW phase will be treated with the same dosage or volume (2 mg/kg, s.c., every 2 days) of either TA or equivalent volume of placebo, as randomized. The treatment period in this phase is up to 16 weeks.

6.3.2 Method of Assigning Patients to Treatment Groups

Patients who meet all inclusion and non-exclusion criteria will be enrolled in the SAOL phase and will be scheduled to receive open-label TA for 18 weeks. All patients who have completed 18 weeks of treatment during the SAOL phase and showing a partial or complete response at Week 18 (V8) will be randomized to TA or placebo in a 1:1 ratio to begin treatment in the RW phase. Experiencing a disease reactivation at any point during the 18-week SAOL phase will not prevent a patient from enrollment into the RW phase; rather, patients are required to meet the criteria for ongoing TA treatment throughout the SAOL phase and be without show a partial or complete response at Week 18. The randomization will be stratified by disease condition at RW phase entry (NLRC4-MAS mutation or XIAP deficiency). The randomization will be designed to be dynamic yet non-deterministic such that patients are randomized in a 1:1 ratio within each disease condition, while also ensuring that the overall allocation is exactly balanced if the total number of patients randomized is an even number and different by no more than one if odd.

Given the study specifications to randomize ten subjects in each of two strata defined by genetic mutation, a randomization strategy is defined based on a binomial distribution which allocates subjects in a 1:1 ratio to TA or Placebo. This will be accomplished by a blocked randomization, wherein within each block of two subjects, one subject is randomized to receive TA and the other is randomized to receive Placebo. With the planned fixed sample size of N=10 and a block size of two, the randomization list will have five blocks. Because there are two possible configurations of each block (TA, Placebo; and Placebo, TA), there are a total of 32 different ways the five blocks can be arranged to produce the randomization list.

6.4 Efficacy and Safety Variables

6.4.1 Efficacy Variables

Efficacy assessment is related to the incidence of disease reactivations and will cover:

i) Clinical evaluation (captured mostly by the mAIDAI):
   - Complete physical examination including: vital signs (systolic and diastolic blood pressure, heart rate, respiratory rate, weight, and body temperature)

- All findings should be documented: joint swelling, hepatomegaly and splenomegaly, presence of skin rash, and ophthalmic control (presence of uveitis). Ophthalmic examination will be performed for assessment of uveitis;
- Symptoms requiring documentation: abdominal pain, diarrhea, nausea, and vomiting
- All clinical components (including clinical findings and symptoms) of the disease will be captured in a binary way as either being present or absent. This binary system is considered the most stringent tool to describe a clinically relevant level of improvement and control of the severe clinical manifestations in the patients treated with TA. The judgement, of whether an end organ category is affected, is done by the treating physician according to clinical standards, as also previously done for all mAIDAI components
- Daily weight, stool output, and PO intake measurements including number of emesis episodes will be performed on patients with intestinal dysfunction, only when the patients are hospitalized
- Radiologic findings if assessed as necessary by the treating physician

ii) Maintenance of co-treatment:

- Total weekly glucocorticoids
- IL-1 blockers
- Other treatments (ie, disease modifying anti-rheumatic drug, etc.)

iii) Laboratory assessments:

- Serum CRP
- Serum ferritin
- Complete blood counts
- Serum levels of total IL-18, free IL-18, and IL-18BP

6.4.1.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the time to first occurrence of disease reactivation (including full and partial disease reactivation) during the 16-week RW phase.

A Disease reactivation occurs if:

For disease reactivation, only the objective mAIDAI components will be evaluated as indicated below. Subjective mAIDAI components (see Appendix I of the clinical study protocol) will not be considered to avoid any bias in the assessment. Appendix

A includes a description of the mapping of the mAIDAI components into the Objective mAIDAI total score.

During the RW-phase, the reference for a disease reactivation (for the primary endpoint assessment) will be the end of the SAOL phase (V8); for the SAOL phase, reference will be the previous visit, as patients enter the SAOL phase with an active disease.

- **Partial disease reactivation**
  - Increase in the number of affected objective end organ damage categories by at least 2 or
  - Increase in the number of affected objective end organ damage categories by at least 1 and worsening of at least 1 systemic inflammatory marker (i.e., serum ferritin levels ⩾500 ng/mL or serum CRP ⩾2x ULN) that was meeting complete response criteria before
- **Full disease reactivation**
  - Increase in the number of affected objective end organ damage categories to at least the number of affected end organ damage categories at baseline (V1) and
  - Exacerbation of at least 1 systemic inflammatory marker (i.e., serum ferritin levels ⩾500 ng/mL or serum CRP ⩾2x ULN) that was meeting complete response criteria before

Note: It is recommended that patients suffering from *severe* CNS manifestations are put on rescue immunosuppression treatment.

Time to first occurrence of disease reactivation will be calculated in weeks as the date of disease reactivation during the RW phase minus the date of randomization during the RW phase plus 1, divided by 7. Patients who do not experience a disease reactivation during the RW phase will be censored at the date of their last clinical assessment during the RW phase indicating no disease reactivation. Patients will be assessed for disease reactivation during the RW phase at all scheduled or unscheduled visits.

6.4.1.2 Secondary Efficacy Endpoints

The key secondary efficacy endpoint is

- Response to therapy (including complete response and partial response) in the SAOL phase from Week 10 onwards and observed at least at 2 consecutive visits at least 2 weeks apart during the SAOL phase

Other secondary efficacy endpoints include the following:

- Best response to therapy during the SAOL phase (including complete response and partial response and disease improvement)

- Duration of response to therapy (including complete response and partial response) during the SAOL phase
- Disease reactivation rate (ie, number of disease reactivations experienced per week while each subject is on the study treatment during the SAOL phase)
- Treatment failures (ie, patients who experience at least one disease reactivation), during the RW phase
- Intensity of disease reactivations (defined by the level of activity given by the mAIDAI – both objective and subjective criteria), during the RW phase and the SAOL phase
- Serum CRP during the RW phase and the SAOL phase
- Serum ferritin during the RW phase and the SAOL phase
- Improvement of fevers (if present at baseline) during the RW phase and the SAOL phase
- Improvement of hepato/splenomegaly (if present at baseline) during the RW phase and the SAOL phase
- Improvement in serum albumin and liver transaminases, anemia and/or platelet count (if present at baseline) during the RW phase and the SAOL phase
- Number of hospitalizations and hospital length of stay during the RW phase and the SAOL phase
- PGA during the RW phase and the SAOL phase
- Patient's or caregiver's qualitative evaluation of health status
- Presence of skin rash (evolution if present at baseline or appearance during the study) during the RW phase and the SAOL phase
- Tolerance to oral/enteral nutrition for hospitalized patients with intestinal dysfunction that was present at baseline during the RW phase and the SAOL phase
- Improvement of stool output for hospitalized patients if intestinal dysfunction was present at baseline during the RW phase and the SAOL phase

Active disease upon enrolment is defined as:

- Increased inflammatory markers, i.e., serum ferritin levels ≥ 500 ng/mL or serum CRP ≥ 2x ULN, and
- Active clinical disease as evidenced by mAIDAI ≥ 4 or CNS manifestations (i.e., seizures, altered mental status, signs of increased intracranial pressure, chronic papilledema, loss of vision, other sensorineural deficiencies, etc.)

For the definition of active disease upon enrolment, all mAIDAI components (subjective and objective) will be evaluated in line with previous protocol versions prior to version 6.0.

Response to Therapy is defined as:

1. In the 16-week RW phase, response to therapy (as primary endpoint of efficacy) will be assessed by prevention of disease reactivations during the 16-week treatment in the RW phase.
2. In the SAOL phase, response to therapy (as secondary endpoint of efficacy) will be assessed through the control of the active disease

For response to therapy, only the objective mAIDAI components will be evaluated as indicated below. Subjective mAIDAI components (see Appendix I of the clinical study protocol) will not be considered to avoid any bias in the efficacy assessment in the open label part of the study.

- **Complete response**

  No major end organ damage (=objective mAIDAI components)

  - No fever (body temperature ≤100.4ºF/38˚C)
  - No diarrhea
  - No transaminitis and/or hepatomegaly
  - No splenomegaly and/or adenopathy
  - No rash
  - No uveitis
  - No arthritis
  - No radiologic CNS involvement
  - No cytopenia (HgB ≥9.0 and/or PLT ≥100 and/or WBC ≥3.0)

  No systemic inflammation

  - No hyperferritinemia (serum level < 500 ng/mL)
  - No increased CRP (serum level < 2x ULN)

  Immunosuppressive therapies

  - Discontinuation of systemic corticosteroids (excluding hydrocortisone for prevention of adrenal insufficiency) or other immunosuppressive therapies including biologics ongoing at baseline
  - No start or dose increase of other immunosuppressive therapies including biologics

- **Partial response**

  At least 50% reduction of the number of affected objective end organ damage categories at baseline and

- At least 50% of the systemic inflammatory markers (ferritin and/or CRP) increased at baseline reaching levels for complete response or
- Discontinuation of systemic corticosteroids (excluding hydrocortisone for prevention of adrenal insufficiency) or other immunosuppressive therapies including biologics ongoing at baseline

Partial response is thus a measure of the complete resolution of clinical manifestations in at least 50% of involved specific organ classes together with a reduction of the systemic inflammatory markers most relevant for this disease or the possibility to reduce other immunosuppressive therapies.

- **Disease Improvement**

  A third response category of Disease Improvement is introduced to be utilized as a description of best response and in order to better describe initial treatment response to therapy and disease evolution over time. This is only used for analysis purposes and will not be evaluated for the decision whether to be eligible for the start of the RW phase. Disease Improvement is defined as:

  - Reduction of the number of affected objective end organ damage categories at baseline by at least 1 or
  - Normalization (i.e., reaching levels for complete response) or improvement (> 50% change from baseline for laboratory markers but still exceeding levels for complete response) of at least 50% of the systemic inflammatory markers affected at baseline
  - Systemic corticosteroids (excluding hydrocortisone for prevention of adrenal insufficiency) or other immunosuppressive therapies including biologics must be at the same or lower doses compared to baseline

6.4.2 Description of Safety and Tolerability Variables

Safety and tolerability variables include the following:

- Adverse events (including adverse events of special interest [AESI])
- Physical examination findings including an ophthalmological examination (ie, clinically significant changes from Baseline)
- Vital signs (ie, clinically significant changes from SAOL Baseline for data collected during the SAOL phase and RW Baseline for data collected during the RW phase)
- Laboratory assessments including clinically significant changes from SAOL Baseline for data collected during the SAOL phase and RW Baseline for data collected during the RW phase in hematology with platelet counts, CRP, ESR, ferritin, fibrinogen, D-dimer, liver enzymes, creatinine, and glomerular filtration rate

- Immunogenicity evaluation: Generation of anti-rh-IL-18BP antibodies
- Local tolerability at the injection site, as evaluated by a standardized assessment

An AESI for this study is defined as an event that constitutes an injection site reaction, including pruritus, erythema, and swelling.

6.4.3 Description of Pharmacokinetic Variables

The following pharmacokinetic (PK) parameters will be estimated at Week 0 (V1) and Week 4 (V5) by non-compartmental methods using actual elapsed time from dosing:

- $AUC_{0\text{-}24}$: Area under the concentration-time curve (ng × h/mL) from time zero (pre-dose) to 24 hours post-treatment, calculated by the linear trapezoidal rule
- C_{max}: Maximum concentration (ng × mL), obtained directly from the observed concentration versus time data
- T_{max}: Time to maximum concentration (hours), obtained directly from the observed concentration versus time data
- $T_{1/2}$: Apparent terminal half-life (hours), determined as $\ln(2)/\lambda_z$
- λ_z: Apparent terminal elimination rate constant (1/hours), determined by linear regression of terminal points of the log-linear concentration-time curve. A minimum of 3 data points are required for determination of λ_z, with an r-square value ≥0.8. In addition, visual assessment by the pharmacokineticist may be used to confirm the terminal linear phase of the concentration-time profile

For calculations using the linear trapezoidal rule, the following formula will be used:

$$AAAAAA = \frac{1}{2}(AA_1 + AA_2)(tt_2 - tt_1)$$

6.5 Data Quality Assurance

Report summaries will be generated using validated Base SAS® software, version 9.4 or higher, on a PC or server-based platform. PK parameters and analyses will also be generated using SAS® software. Additional validated software may be used to generate analyses, as needed.

All SAS programs that create outputs or supporting analysis datasets will be validated by a second statistical programmer or biostatistician. At a minimum, validation of programs will consist of a review of the program log, review of output or dataset format and structure, and independent confirmatory programming to verify output results or dataset content. Additionally, all outputs will undergo a review by a senior level team member before finalization.

The content of the source data will be reviewed on an ongoing basis by project statistical programmers and statisticians. Data will be checked for missing values, invalid records, and extreme outliers through defensive programming applications, analysis-based edit checks, and other programmatic testing procedures. All findings will be forwarded to the project data manager for appropriate action and resolution.

7 STATISTICAL METHODS

7.1 General Methodology

Data will be analyzed by Precision for Medicine biostatistics personnel. Statistical analyses will be reported with tables, figures, and listings, presented in rich text format, and using recommended ICH numbering. Output specifications for all tables, figures, and listings will be in conformance with guidelines specified by the ICH in Appendix 7 of the *Electronic Common Technical Document Specification* (Apr 2003).

7.1.1 Reporting Conventions

Tables and figures will be summarized for all patients combined during the SAOL phase and by treatment group during the RW phase. Tables summarizing demographics and other baseline characteristics for the RW phase will also include a column for all patients combined. Where applicable, separate summaries will be presented for each study phase (SAOL and RW). In general, all data collected and any derived data will be presented in patient data listings, for all enrolled patients. Listings will be ordered by site, patient number, treatment group, and assessment or event date. The treatment group presented in listings will be based on the planned assignment, unless otherwise noted.

In general, continuous variables will be summarized to indicate the study population sample size (N), number of patients with available data (n), mean, SD, median, 25th (Q1) and 75th (Q3) quartiles, minimum, and maximum values. Selected continuous efficacy variables will also be summarized by the 95% CI of the mean. Categorical variables will be summarized by the population size (N), number of patients with available data (n), number of patients in each category, and the percentage of patients in each category. Unless otherwise noted, the denominator to determine the percentage of patients in each category will be based on the number of patients with available data. Selected ordinal data may be summarized using both descriptive statistics and counts and percentages of patients in each category, as appropriate.

Non-zero percentages will be rounded to one decimal place. Rounding conventions for presentation of summary statistics will be based on the precision of the variable of summarization, as it is collected in its rawest form (ie, on the eCRF or as provided within an external file) and are outlined as follows:

- The mean and median will be rounded to one more decimal place than the precision of the variable of summarization

- Measures of variability (eg, SD, SE) will be rounded to two more decimal places than the precision of the variable of summarization

- Minimum and maximum values will be presented using the same precision as the variable of summarization

Other statistics (eg, CIs) will be presented using the same general rules outlined above or assessed for the most appropriate presentation based on the underlying data.

Statistical significance testing will be two-sided and performed using α=0.05. P-values will be reported for all statistical tests, rounded to four decimal places. P-values less than 0.0001 will be displayed as "<0.0001"; p-values greater than 0.9999 will be displayed as ">0.9999".

7.1.2 Definition of Baseline

Changes from baseline, percentage changes from baseline, and occurrence of disease reactivations will be calculated separately within each study phase (SAOL and RW), relative to the start of dosing during that phase:

- **SAOL Baseline**: Baseline during the SAOL phase will be defined as the last measurement reported prior to the first dose of open-label TA at Week 0 (V1)

- **RW Baseline**: Baseline during the RW phase will be defined as the last measurement reported prior to the first dose of double-blind study drug (TA or placebo) at Week 18 (V8)

7.1.3 Summarization by Visit

Data summarized by study visit will be based on the nominal, scheduled visit label as reported on the eCRF, except for data that are captured at early termination or unscheduled visits. Data collected at early termination or unscheduled visits will only be included in by-visit summaries if they fall within a scheduled post-baseline visit window (see Table 2) and there is no data from a nominal visit for that specific visit. If both an unscheduled visit and an early termination visit fall within the same visit window the early termination visit will be used for that specific visit.

Table 2 Visit Windows for Early Termination and Unscheduled Visits

Phase	Analysis Visit	Target Day	Visit Window
SAOL	Week 1	SAOL Day 7	2 ≤ SAOL Day ≤ 11
	Week 2	SAOL Day 14	12 ≤ SAOL Day ≤ 18
	Week 3	SAOL Day 21	19 ≤ SAOL Day ≤ 25
	Week 4	SAOL Day 28	26 ≤ SAOL Day ≤ 42
	Week 8	SAOL Day 56	43 ≤ SAOL Day ≤ 70
	Week 12	SAOL Day 84	71 ≤ SAOL Day ≤ 105
	Week 18	SAOL Day 126	106 ≤ SAOL Day ≤ End of SAOL Phase
	Week 20	RW Day 14	2 ≤ RW Day ≤ 21
	Week 22	RW Day 28	22 ≤ RW Day ≤ 42
	Week 26	RW Day 56	43 ≤ RW Day ≤ 70
	Week 30	RW Day 84	71 ≤ RW Day ≤ 98
	Week 34	RW Day 112	99 ≤ RW Day ≤ End of RW Phase

Note: SAOL Day is calculated as visit date – first SAOL dose date +1. RW Day is calculated as visit date – first RW dose date +1.

Data collected at all visits will be considered when endpoint derivations potentially include multiple visits (eg, determination of baseline value, determination worst post-baseline value, etc.). All data will be included in patient listings.

7.1.4 Data Handling Rules

Unless otherwise noted, values reported as greater than or less than some quantifiable limit (eg, "< 1.0") will be summarized with the sign suppressed in summary tables and figures, using the numeric value reported. Data will display on patient listings to include the sign.

In some cases, CRP values from the central laboratory were entered as '< 3 mg/L'. Since the ULN for CRP is 1 mg/L, a value of < 3 mg/L does not allow determination of whether the CRP value is ≥ 2x ULN. In these cases, a re-test of high sensitivity CRP will be attempted, and those values will be used to replace the < 3 mg/L value. If the high sensitivity CRP value is not obtained a local lab CRP value may be used to determine whether the CRP value is ≥ 2x ULN.

7.1.5 Standard Calculations

Where appropriate, the calculated study day of each assessment or event will be presented with the assessment or event date on patient data listings. The calculated study day will be relative to the start of dosing within each phase:

- The assessment/event date minus the date of first dose received during the SAOL or RW phase, if the assessment/event date is prior to the date of first dose; and

- The assessment/event date minus the date of first dose received during the SAOL or RW phase, plus one, if the assessment/event date is on or after the date of first dose.

Other variables requiring calculations will be derived using the following formulas:

- **Days:** A duration between two dates expressed in days will be calculated using the following conventions:
    - Later date – earlier date + 1, if the earlier date is on or after the date of first dose of study drug during the SAOL or RW phase; or
    - Later date – earlier date, if the earlier date is prior to the date of first dose of study drug during the SAOL or RW phase.
- **Weeks:** A duration expressed in weeks will be calculated by the dividing the duration in days by seven
- **Months:** A duration expressed in months will be calculated by dividing the duration in days by (365.25 / 12)
- **Years:** A duration expressed in years will be calculated by dividing the duration in days by 365.25
- **Change from Baseline:** Change from baseline will be calculated as the post-baseline value minus the baseline value
- **Percentage Change from Baseline:** Percentage change from baseline will be calculated as the change from baseline divided by the baseline value, multiplied by 100

7.1.6 Patient Consideration for Original Study Design

Patients enrolled under the original protocol through Protocol Amendment 2 (Version 3.0) and who were randomized to TA during the RDBPC phase of the study will be considered for all analyses, combined with those patients enrolled under Protocol Amendment 4 (Version 4.0), or any subsequent amendments. Patients randomized to placebo under the original study design will only be included in patient data listings and will be characterized separately in the CSR.

7.2 Analysis Populations

The analysis populations are defined as follows:

- Single Arm Open-Label Full Analysis Set (SAOL-FAS): All patients who are eligible and receive at least one dose of open-label TA at Week 0 (V1) under Amendment 3 (Version 4.0) of the protocol (or subsequent amendments), or at

least one dose of TA at Week 0 (V1) during the RDBPC phase under previous protocol amendments.

- Single Arm Open-Label PK Analysis Set (SAOL-PKAS): All patients in the SAOL-FAS who have sufficient plasma concentration data to calculate PK parameters. If any patients are found to be noncompliant with respect to dosing or have incomplete data, a decision will be made prior to DBL on a case-by-case basis as to their inclusion in the analysis.

- Randomized Withdrawal Full Analysis Set (RW-FAS): All patients who meet the criteria for randomization and are randomized to blinded study treatment during the RW phase.

- Randomized Withdrawal Per-Protocol Analysis Set (RW-PPAS): All patients in the RW-FAS who complete the RW phase up through Week 26 (V11) for protocol version 4.0/4.1 or Week 34 (V13) for protocol versions 5.0 and 6.0 or have disease reactivation and for whom no relevant protocol deviations (ie, major protocol deviations) were documented.

The analyses of safety and efficacy during the SAOL phase will be based on the SAOL-FAS and the analyses of safety and efficacy during the RW phase will be based on the RW-FAS. Additional exploratory analyses of efficacy during the RW phase will be performed based on the RW-PPAS. Analysis of PK plasma concentration data and PK parameters will be performed based on the SAOL-PKAS.

7.3 Study Patients

7.3.1 Disposition of Patients

Patient disposition will be presented separately for the SAOL phase and RW phase.

Summaries during the SAOL phase will include the number and percentage of patients in each analysis population, completing TA treatment during the SAOL phase, discontinuing TA early during the SAOL phase by reason for discontinuation, starting new medications for treatment of the patient disease, completing the SAOL phase, and discontinuing the SAOL phase early by the primary reason for discontinuation.

Summaries during the RW phase will be based on the RW-FAS and will be summarized by randomized treatment received during the RW phase. The summary will include the number and percentage of patients completing study treatment during the RW phase, discontinuing study treatment during the RW phase by reason for discontinuation, starting new medications for treatment of the patient disease, completing the RW phase, and discontinuing the RW phase early by primary reason for discontinuation.

7.3.2 Protocol Deviations

Major protocol deviations occurring during the RW phase will be summarized by treatment group and over all patients combined for the RW-FAS. Major protocol

deviations may include, but will not be limited to violation of eligibility criteria, randomization errors, lack of compliance with study drug dosing, or on-study administration of a prohibited medication.

All major protocol deviations will be determined and appropriately categorized prior to database lock and prior to breaking the blind of the treatment group assignments. The decision as to whether a protocol deviation is a major protocol deviation or a major protocol violation leading to exclusion from the RW-PPAS Analysis Set will be made on a case-by-case basis in a blinded data review meeting. The number and percentage of patients with any major protocol deviations as well as the number and percentage of patients with major deviations within each category will be presented.

All protocol deviations collected on study during either the SAOL phase or RW phase will be listed by patient to include the date the deviation occurred, type of deviation, and a description.

7.3.3 Demographic and Other Baseline Characteristics

Demographic variables including age, sex, ethnicity, and race will be summarized for the SAOL-FAS, SAOL-PKAS, RW-FAS, and RW-PPAS. Age as reported within the electronic data capture system will be summarized using descriptive statistics. Sex, ethnicity, and race will be summarized with the number and percentage of patients in each parameter category.

Baseline characteristics include disease history, other manifestations of intestinal dysfunction, other clinical symptoms of disease, medical history, disease-specific prior medication use, height, weight, and body mass index (BMI).

Disease history, other manifestations of intestinal dysfunction, and other clinical symptoms of disease will be summarized for the SAOL-FAS, RW-FAS, and RW-PPAS by treatment group and over all patients combined.

Height, weight, BMI, Medical history, and prior medication use will be summarized for the SAOL-FAS.

Height, weight, and BMI at baseline will be summarized using descriptive statistics. Frequency counts and percentages to summarize patients reporting abnormal medical history by body system will be presented.

Baseline disease history will be described by the time elapsed (months) since diagnosis of disease and the incidence of previous and current disease history by symptom. Time since diagnosis of disease will be calculated as described in Section 7.1.4 where complete diagnosis dates are provided. If only the month and year are provided, the months since diagnosis will be determined as difference in months between the diagnosis month and year and the month and year of informed consent. If only the year of diagnosis is provided, the months since diagnosis will be determined as the difference between the diagnosis year and the year of informed consent, multiplied by 12. Months since diagnosis will be summarized using descriptive statistics. The number and percentage of

patients reporting each disease history symptom, including other manifestations of intestinal dysfunction and other clinical symptoms and/or signs, will be presented separately for previous history and current history.

Disease-specific prior medications will include those treatments specific to the treatment of the auto-inflammatory condition that started and stopped prior to the date the informed consent is signed, as reported on the Prior Medications eCRF. Medications will be coded using the World Health Organization Drug Dictionary Enhanced (WHODDE, version March 1, 2017). Medications entered on the eCRF will be mapped to Anatomic Therapeutic Chemical (ATC) drug class (level 4) and drug name. The number and percentage of patients receiving any prior medication will be summarized by treatment group and over all patients combined, as will the number and percentage receiving any medication by ATC drug class and generic drug name. Patients reporting use of more than one medication at each level of summarization (any prior medication received, ATC class, and generic drug name) will be counted only once. ATC class terms will be displayed by descending order of incidence, as will generic drug names within each ATC class.

7.3.4 Measurements of Treatment Compliance

Compliance to the study treatment regimen will be determined as the total number of days with a dose received divided by the expected number of days with a dose received, multiplied by 100. The total number of days with a dose received will be determined by phase based on entries on the Study Drug Administration form, where a dose received is counted for a given day if the response to "Was the study drug administered" is "Yes." The expected number of days with a dose received will be calculated as the date of last dose of study drug minus the date of first dose of study drug plus one, divided by two. Calculations that result in a value with a remainder of one half (ie, the number of days between doses is an odd number) will be rounded up to the next largest integer. The date of last dose of study drug is the date reported on the End of Treatment form, for both the end of double-blind treatment and the end of treatment during the RW phase.

Dosing compliance will be summarized using descriptive statistics, by treatment group, based on the SAOL-FAS for dosing during the SAOL phase and based on the RW-FAS for dosing during the RW phase. The number and percentages of patients who are $< 80\%$ compliant and $\geq 80\%$ compliant within each treatment group will be summarized.

7.4 Efficacy Evaluation

7.4.1 Datasets Analyzed

All efficacy summaries for data collected during the RW phase will be based on the RW-FAS; selected efficacy summaries will also be produced on the RW-PPAS provided the RW-FAS and RW-PPAS differ by at least one patient. A data listing of patients excluded from the RW-FAS or RW-PPAS, including the reason for exclusion, will be presented.

All efficacy summaries of data collected during the SAOL phase will be based on the SAOL-FAS.

7.4.2 Primary Efficacy Endpoint Analysis Methods

The primary efficacy endpoint is the time to first occurrence of disease reactivation (including full and partial disease reactivation) during the 16-week RW phase, as defined in Section 6.4.1.1.

7.4.2.1 Primary Estimand of the Primary Efficacy Endpoint

The primary estimand for the primary endpoint will be described as follows:

Population: All patients in the RW-FAS population.

Endpoint: Time to first occurrence of disease reactivation during the RW phase.

Intercurrent events:

- End of RW phase: patients who do not experience a disease reactivation and complete the RW phase will be censored at the time of their last disease assessment during the RW phase regardless of the length of the RW phase (according to the protocol version the patient was treated under).
- Rescue immunosuppression treatment: patients who receive rescue immunosuppression treatment without objective disease reactivation and continue in the RW phase will be censored at the time of their last disease assessment during the RW phase, including assessments up to the start day of the rescue immunosuppression treatment.
- Clinical signs of inflammation that are not disease related: patients who present with clinical or laboratory signs of inflammation meeting the criteria of disease reactivation that are determined by the treating physician to be related to an infection rather than disease reactivation and are resolved after repeat assessments within the visit window will not be considered to have a disease reactivation at that visit (see Section 7.4.2.2).

Population-level summary: Kaplan-Meier (KM) estimates of the survival curves by treatment group, hazard ratio including 95% CI, and log-rank test.

The primary efficacy analysis will test the following hypothesis:

- H_0: The survival distribution of time to first occurrence of disease reactivation during the RW phase is equal between TA and placebo
- H_1: The survival distribution of time to first occurrence of disease reactivation during the RW phase is different between TA and placebo

TA will be compared to placebo using the log rank test. Kaplan-Meier (KM) estimates of the distribution of time-to-event will be summarized and graphed by treatment group. Greenwood's formula will be used to estimate the standard error of the KM statistic. The

primary analysis endpoint will be tested at a 2-sided alpha level of 0.05. The summarization will include the KM estimate of the median, 25th and 75th quartiles, and corresponding 95% CI, if estimable. Summaries will include the "at-risk patient" counts at the scheduled post-baseline visits during the RW phase. The number and percent of patients experiencing a disease reactivation and censored will be presented. The hazard ratio and 95% CI will be determined based on a Cox regression model with a single factor for treatment group.

7.4.2.2 Primary Efficacy Endpoint evaluation in case of repeat assessments

The primary endpoint is defined as the time to first occurrence of disease reactivation during the RW phase. To distinguish between a disease reactivation and an appearance of clinical signs of an infection, the treating physician may decide to repeat clinical assessments within the protocol defined visit window. Those additional assessments are documented as a separate unscheduled visit in the database.

The handling of such repeat assessments is currently not described in the study protocol. If such repeat assessments are available, the evaluation of the treating physician whether clinical or laboratory signs of inflammation are considered as transient, infection-related manifestations or signs of a disease reactivation should be considered for the determination of a disease reactivation per protocol. This evaluation is done in a blinded manner prior to database lock and documented accordingly.

7.4.2.3 Sensitivity Analyses of the Primary Efficacy Endpoint

The primary endpoint analysis will be repeated on the RW-PPAS, provided the RW-FAS and RW-PPAS differ by at least one patient.

In addition, the following tipping point analyses will be performed for the primary endpoint to assess the impact of the different lengths of the RW phase due to the different versions of the protocol. For patients enrolled in protocol versions 4.0/4.1 and earlier whose planned RW duration was only 8 weeks and had no disease reactivation during the RW phase:

1. A worst-case imputation will be applied to patients in both treatment groups by assigning patients as having disease reactivation at the next day after the last per protocol visit.

2. A best-case imputation will be applied to patients in both treatment groups by assigning patients as not having disease reactivation until Week 16 of the RW phase.

3. A worst-case imputation will be applied to patients in the TA treatment group and a best-case imputation will be applied to patients in the Placebo treatment group.

7.4.3 Key Secondary Efficacy Endpoint Analysis Methods

The key secondary efficacy endpoint is response to therapy (including complete response and partial response) in the SAOL phase from Week 10 onwards and observed at least 2 consecutive visits at least 2 weeks apart. Patients who do not achieve complete response or partial response at 2 consecutive visits at least 2 weeks apart at Week 10 onwards or drop out prior to Week 10 will be considered non-responders. Patients who respond to therapy in the SAOL phase with complete or partial response will be summarized by the number and percent of patients along with the associated 95% exact Clopper-Pearson CIs.

The estimand for this endpoint will be described as:

Population: All patients in the SAOL-FAS population

Endpoint: Responder or non-responder.

Intercurrent events:

- Discontinuation prior to Week 10: patients who discontinue prior to Week 10 will be assigned as non-responders.
- Rescue immunosuppression treatment: patients who receive rescue immunosuppression treatment and continue in the SAOL phase and satisfy the criteria of responder will be considered responders.

Population-level summary: Number and percentage of responders and 95% exact Clopper-Pearson CIs.

A sensitivity analysis will also be performed that only considers assessments up to the start day of the rescue immunosuppression treatment.

7.4.4 Secondary and Additional Efficacy Endpoint Analysis Methods

Secondary endpoints collected during the RW phase will be summarized for the RW-FAS. Selected secondary endpoints will also be summarized for the RW-PPAS, provided the RW-FAS and RW-PPAS differ by at least one patient. Reference Section 7.4.6.5 for detail on testing procedures for secondary endpoints (ie, multiplicity adjustments).

Secondary endpoints collected during the SAOL phase will be summarized for the SAOL-FAS.

7.4.4.1 Best Response to Therapy During SAOL

The number and percent of subjects in each response category will be summarized, including: complete response, partial response, disease improvement, and no improvement or not evaluated.

7.4.4.2 Treatment Failures during RW

Patients meeting the criteria for treatment failure during the RW phase will be summarized by the number and percent of patients in each treatment group. The estimated difference in percentages (TA – placebo) among treatment failures between treatment groups will be presented along with the associated 95% exact CI. The 95% CIs for difference in binomial proportions will be computed based on exact unconditional confidence limits using the score statistic. Treatment groups will be compared for response to therapy versus treatment failure using the Fisher's exact test.

The estimand for this secondary endpoint will be described as follows:

Population: All patients in the RW-FAS population.

Endpoint: Treatment failure (Yes/No).

Intercurrent events: All patients who experience disease reactivation during the RW phase will be counted as treatment failure regardless of intercurrent events.

Population-level summary: Number and percentage of treatment failures, exact unconditional 95% CIs using the score statistic, and Fisher's exact test.

7.4.4.3 Change from Baseline in C-reactive Protein and Serum Ferritin

Descriptive statistics for the observed CRP and serum ferritin values and the changes from RW Baseline during the RW phase will be presented by visit for each treatment group. Descriptive statistics will include number of subjects, mean, standard deviation, 95% CI of the mean, median, Q1, Q3, minimum, and maximum values. Mean CRP and serum ferritin values with associated standard error bars will be plotted by treatment group over time during the RW phase. The primary comparison of change from RW Baseline values among treatment groups will be based on the Mann-Whitney-Wilcoxon test. As sensitivity analyses, the mean change from RW baseline will also be compared among treatment groups using Welch's two-sample test that allows for unequal variances, as well as by a linear regression model with Huber-white sandwich errors allowing for heteroscedasticity.

To assess the impact of possible variability in RW Baseline measures between groups, a rank analysis of covariance (ANCOVA) model with covariate adjustment for RW Baseline will be performed as a sensitivity analysis. This will be performed at all post-baseline visits during the RW phase that include a sufficient number of patients within each treatment group to allow for appropriate model testing. Model estimates for the linear regression model and the rank ANCOVA model will be presented to include the estimate of mean treatment difference, associated standard error, 95% CI of the mean difference, and p-value.

Descriptive statistics for the observed CRP and serum ferritin values and changes from baseline during the SAOL phase will be presented by visit.

7.4.4.4 Modified Auto-Inflammatory Disease Activity Index

Descriptive statistics for the mAIDAI total score (including both objective and subjective components) and changes from RW Baseline during the RW phase will be presented by visit for each treatment group. Descriptive statistics will include number of subjects, mean, standard deviation, 95% CI of the mean, median, Q1, Q3, minimum, and maximum values. Mean mAIDAI total score with associated standard error bars will be plotted by treatment group over time during the RW phase.

The primary comparison for the change from baseline in the mAIDAI total score among treatment groups will be based on the Mann-Whitney-Wilcoxon test.

The estimand for the secondary endpoint of change in mAIDAI total score from RW baseline to Week 34/ET will be described as follows:

Population: All patients in the RW-FAS population.

Endpoint: Change from RW baseline to Week 34.

Intercurrent events:

- Discontinuation prior to Week 34: change from baseline at Week 34 for patients who discontinue prior to Week 34 or for patients who were enrolled under protocol v4.0/4.1 and earlier and complete RW at Week 26 will be imputed using last observation carried forward (LOCF).
- Rescue immunosuppression treatment: patients who receive rescue immunosuppression treatment and continue in the RW phase will include all data captured within the RW phase including visits after receiving rescue immunosuppression treatment.

Population-level summary: Descriptive statistics and Mann-Whitney-Wilcoxon test.

As sensitivity analyses, change from baseline to maximum post-baseline score and change from baseline to last RW assessment up to the start day of rescue immunosuppression treatment will be summarized and compared among treatment groups. In addition, the mean change from RW baseline in the mAIDAI total score will also be compared among treatment groups using Welch's two-sample t-test that allows for unequal variances, as well as by a linear regression model with Huber-white sandwich erros allowing for heteroscedasticity. A rank ANCOVA model to include a covariate for the RW Baseline score will also be performed as sensitivity analysis, similar to the analysis of CRP and serum ferritin changes from RW Baseline. Model estimates for the linear regression model and the rank ANCOVA model will be presented to include the estimate of mean treatment difference, associated standard error, 95% CI of the mean difference, and p-value.

Shifts in mAIDAI responses from RW Baseline to each post-baseline visit during the RW phase will be summarized by component and treatment group (ie, present or absent at

baseline to present or absent at each post-baseline visit). Summarization will include the total counts for all responses reported at both RW Baseline and each post-baseline visit. Percentages will be based on the number of patients with a non-missing RW Baseline and non-missing post-baseline assessment at each visit.

The number and percent of patients who improve within each component (objective and subjective) of the mAIDAI will be summarized during the RW phase by visit and treatment group. For each component, only the number of patients who reported the assessment as present (1 or 2) will be considered for analysis. The proportion of patients who report the assessment as absent (0, ie, improved relative to baseline) at each post-baseline visit during the RW phase will be compared among treatment groups using the Fisher's exact test. Patients with a missing assessment at a post-baseline visit will be considered to have not improved.

Descriptive statistics for the mAIDAI total score and changes from baseline during the SAOL phase will be presented by visit. Shifts in mAIDAI responses from baseline during the SAOL phase will be presented by visit. The number and percent of patients who improve within each component of the mAIDAI during the SAOL phase will be presented by visit.

7.4.4.5 Physician's Global Assessment

Descriptive statistics for the PGA symptom severity score and changes from RW Baseline during the RW phase will be presented by visit for each treatment group. Descriptive statistics will include number of subjects, mean, standard deviation, 95% CI of the mean, median, Q1, Q3, minimum, and maximum values. Mean PGA symptom scores with associated standard error bars will be plotted by treatment group over time during the RW phase.

The primary comparison of the change from RW Baseline in the PGA symptom severity score among treatment groups will be based on the Mann-Whitney-Wilcoxon test.

The estimand for the secondary endpoint of change in PGA symptom severity score from RW baseline to Week 34/ET will be described as follows:

Population: All patients in the RW-FAS population.

Endpoint: Change from RW baseline to Week 34.

Intercurrent events:

- Discontinuation prior to Week 34: change from baseline at Week 34 for patients who discontinue prior to Week 34 or for patients who were enrolled under protocol v4.0/4.1 and earlier and complete RW at Week 26 will be imputed using LOCF.

- Rescue immunosuppression treatment: patients who receive rescue immunosuppression treatment and continue in the RW phase will include all

data captured within the RW phase including visits after receiving rescue immunosuppression treatment.

Population-level summary: Descriptive statistics and Mann-Whitney-Wilcoxon test.

As sensitivity analyses, change from RW baseline to maximum post-baseline score and change from baseline to last RW assessment up to the start day of rescue immunosuppression treatment will be summarized and compared among treatment groups. In addition, the mean change from RW baseline in PGA symptom severity score will also be compared among treatment groups using Welch's two-sample t-test that allows for unequal variances, as well as by a linear regression model with Huber-white sandwich erros allowing for heteroscedasticity. A rank ANCOVA model to include a covariate for the RW Baseline score will also be performed as sensitivity analysis, similar to the analysis of CRP and serum ferritin changes from RW Baseline. Model estimates for the linear regression model and the rank ANCOVA model will be presented to include the estimate of mean treatment difference, associated standard error, 95% CI of the mean difference, and p-value.

Shifts in PGA for stability of the patient's disease-related (auto-inflammatory) symptoms over the previous two weeks at RW Baseline to each post-baseline visit will be summarized by group (ie, improved, stable or worse at baseline to improved, stable, or worse at each post-baseline visit). Summarization will include the total counts for all responses reported at both baseline and each post-baseline visit. Percentages will be based on the number of patients with a non-missing RW Baseline and non-missing post-baseline assessment at each visit.

Descriptive statistics for the PGA symptom severity score and changes from baseline during the SAOL phase will be presented by visit. Shifts in PGA for stability of the patient's disease-related (auto-inflammatory) symptoms during the SAOL phase will be presented by visit.

7.4.4.6 Patient's or Caregiver's Qualitive Evaluation of Health Status

Descriptive statistics for the overall disease-related symptoms score ranging from 0 (none) to 10 (extreme) of the patient/caregiver qualitative evaluation of health status questionnaire and changes from RW Baseline to end of RW phase will be presented for each treatment group. The total domain scores (sum of individual symptom scores) for general wellbeing, gastrointestinal symptoms, musculoskeletal symptoms, skin symptoms, eye symptoms, central nervous system, and lymphatic system will be summarized similarly. The total domain scores will be calculated as the sum each individual symptom score, ranging from 0 (none) to 5 (very severe), within the correspoinding domain. Descriptive statistics will include number of subjects, mean, standard deviation, 95% CI of the mean, median, Q1, Q3, minimum, and maximum values.

The patient/caregiver evaluation of change in symptoms from RW Baseline to end of RW phase will also be summarized by treatment group. The evaluation of the stability of

disease-related symptoms from RW Baseline to end of RW phase will be scored as worse, slightly worse, stable, slightly improved, or improved. Summarization will include the number and percent of patients in each category.

Summaries of the patient/caregiver qualitative evaluation of health status during the SAOL phase will also be presented. Descriptive statistics of means and mean changes from SAOL Baseline to end of SAOL phase for the overall disease-related symptoms score and each total domain score will be summarized. The evaluation of the stability of disease-related symptoms from SAOL Baseline to end of SAOL phase will also be presented.

7.4.4.7 Hospitalization

The number and percent of patients with no hospitalization, 1 hospitalization, 2 hospitalizations, and ≥ 3 hospitalizations occurring during the RW phase will be presented by treatment group, among those patients who were not hospitalized during the SAOL Baseline period. Descriptive statistics for the total hospital length of stay in days will also be presented by treatment group, where the total length of stay for each patient is the sum of the durations for all individual hospitalization admissions. The duration of an individual hospitalization is determined as the date of discharge minus the date of admission, plus one.

Improvement in oral/enteral nutrition and stool output at hospitalizations occurring during the RW phase will be assessed for those patients that report intestinal dysfunction during a baseline hospitalization prior to the first dose of study drug during the SAOL phase. Specifically, descriptive statistics to include changes from the baseline hospitalization will be presented by treatment group for PO intake (mL), number of emesis episodes, and stool output (g) for each post-baseline hospitalization experienced in sequence.

Similar summaries will be presented for hospitalizations and improvement in oral/enteral nutrition and stool output at hospitalizations occurring during the SAOL phase.

7.4.4.8 Cytokines and Inflammatory Mediators

Descriptive statistics for the additional efficacy endpoint analysis of observed serum levels of total IL-18, free IL-18, and IL-18BP and changes from baseline during the RW phase will be presented by visit for each treatment group. A similar summary will be presented for the SAOL phase.

7.4.4.9 Duration of Response during SAOL

Duration of response during the SAOL phase will be summarized using KM estimates for patients who achieve either complete response or partial response during SAOL. KM estimates will include estimates of quartiles (25^{th} percentile, median, 75^{th} percentile) along with corresponding 95% CIs. Duration of response is defined as the time from first assessment indicating partial or complete response until the time of subsequent disease reactivation. Patients who do not experience a disease reactivation during SAOL

following response to therapy will be censored at the date of their last clinical assessment during the SAOL phase.

7.4.4.10 Disease Reactivation Rate per Week during SAOL

The number of disease reactivations experienced during the SAOL phase will be summarized with counts and percentages, for those experiencing 0, 1, 2, 3 or ≥ 4 disease reactivations. The number of disease reactivations per week will be summarized using descriptive statistics. The total number of patient weeks on study treatment during the SAOL phase and the total number of disease reactivations experienced will be presented.

7.4.5 Statistical/Analytical Issues

7.4.5.1 Adjustments for Covariates

Due to the small patient sample size, the changes from RW Baseline for continuous efficacy endpoints will be analyzed using the non-parametric Mann-Whitney-Wilcoxon test, with no adjustments for covariates. However, to assess possible variability in RW Baseline measures between groups, a rank ANCOVA model to include a covariate for the RW Baseline value will be performed as a sensitivity analysis.

7.4.5.2 Handling of Dropouts or Missing Data

The analysis of the primary endpoint, time to first occurrence of disease reactivation, will be based on KM survival analysis techniques, where patients who do not experience a disease reactivation during the RW phase will be censored at date of their last assessment for disease symptoms.

Analysis of response to therapy and treatment failures will consider those patients who have missing relevant data for evaluation of response to be non-responders or treatment failures. Analysis of the improvement in the individual components of the mAIDAI will consider those patients with a given symptom reported as present at baseline while missing the assessment at a post-baseline visit to have not improved at that visit.

All analyses of changes from baseline in continuous measures by visit to include CRP, serum ferritin, mAIDAI total score, and the PGA severity score will be based on observed data. This will include an analysis at Week 26 / 34 / ET to capture the last observed data assessed for patients while on study treatment during the RW phase (ie, combines data collected at Week 26 [V11] or Week 34 [V13] for patients who complete the treatment period as planned with data collected at the Study Completion [V11 / V13] visit for patient who discontinue the RW phase early). All data will be included in the statistical analyses regardless of adherence to study treatment or use of ancillary medication.

Endpoints summarized as a shift from baseline to include mAIDAI responses by component and the PGA for stability of patient's disease related symptoms will be based on observed data, where the denominator to compute the percentage of patients for each

shift category will be based on the number of patients with relevant non-missing baseline and post-baseline at the visit of interest.

For all patients enrolled prior to protocol v5.0, the organomegaly subtypes were not explicitly documented with the mAIDAI. Thus, for each visit where organomegaly is indicated as present but no eCRF page of "subtypes of organomegaly" is completed, the subtypes will be evaluated by corresponding entries in the disease history, physical examination, or adverse events. However, as the physical examination after the baseline visit only captures new or worsened conditions, an improvement of one or multiple organomegaly subtypes may not be documented. Thus, all organomegaly subtypes will be documented as ongoing as long as organomegaly is indicated as being present.

7.4.5.3 Interim Analyses and Data Monitoring

There are no interim analyses planned for this study. A Data Safety Monitoring Board (DSMB) will convene regularly through the duration of the trial to assess the global safety of TA and advise whether to continue or discontinue the study. Complete details of the DSMB roles and responsibilities are outlined in a separate DSMB charter.

7.4.5.4 Multicenter Studies

This is a multicenter study, with approximately eleven clinical sites in the United States, Canada, and Germany. Efficacy data collected from all study centers will be pooled for data analysis. The effect of study center on the efficacy analysis results may be explored post-hoc, as needed.

7.4.5.5 Multiple Comparisons/Multiplicity

The overall 2-sided level of significance will be alpha = 0.05 for all statistical testing to include the primary endpoint analysis. The hypothesis testing of selected secondary endpoints collected during the RW phase will be conducted using a gatekeeping procedure based on a closed fixed-sequence test, provided the primary endpoint comparison is statistically significant ($p < 0.05$). This procedure controls the study-wise type I error as described below. In order to evaluate the comparison for each step, all preceding comparisons must have been statistically significant based on a 2-sided alpha = 0.05 in favor of TA.

1. Placebo and TA will be compared with respect to the primary efficacy endpoint, time to first occurrence of disease reactivation experienced during the RW phase

2. Placebo and TA will be compared for the proportion of treatment failures (ie, experience at least one disease reactivation) during the RW phase

3. Placebo and TA will be compared for results of the change from RW Baseline to Week 34/ET in the mAIDAI total score

4. Placebo and TA will be compared for results of the change from RW Baseline to Week 34/ET in the PGA symptom severity score

If the comparison is not statistically significant at any step defined above using the 2-sided 0.05 alpha level, then remaining comparisons in the stated hierarchy will be considered nominal, descriptive, and exploratory. The study-wise type I error will be maintained with the above closed procedure. All other statistical testing will be evaluated descriptively.

7.4.5.6 Use of an "Efficacy Subset" of Patients

The primary efficacy analysis will be performed on the RW-FAS. Secondary endpoints will also be analyzed for the RW-PPAS provided the RW-FAS and RW-PPAS differ by at least one patient.

7.4.5.7 Active-Control Studies Intended to Show Equivalence

This study does not include an active-control product and is not intended to demonstrate equivalence between any two drug products.

7.4.5.8 Examination of Subgroups

Improvement in oral/enteral nutrition and stool output at hospitalizations occurring will be assessed for the subset of patients that report intestinal dysfunction during a baseline hospitalization.

Beyond that, there are no pre-planned analyses to assess efficacy results by subgroup since the study population is too small to warrant any meaningful interpretations. Select endpoints may be analyzed separately by underlying disease status (ie, NLRC4-MAS, XIAP) post-hoc, as warranted.

7.4.6 Serum Concentrations

Raw serum concentration values will be summarized for the SAOL-PKAS by scheduled sampling time point using descriptive statistics, to include the geometric mean and CV (%). The geometric CV is calculated as 100*sqrt[exp(σ^2)-1], where σ^2 is the variance of the log-transformed data. Each patient's plasma concentrations will be plotted over time. For the purposes of plotting the data, plasma concentrations below the limit of quantification (BLQ) that are imbedded between two measurable concentrations will be set to missing; however, BLQ values occurring prior to the first measurable concentration or after the last measurable concentration will be set to zero. For summaries of plasma concentrations, BLQ values will be set to missing. The number and percentage of patients with BLQ values will be summarized by time point. Mean IL-18BP concentration values with associated standard error bars will be plotted over time during the SAOL phase.

7.4.7 Pharmacokinetic Analysis

Pharmacokinetic parameters will be analyzed based on the SAOL-PKAS. For each patient, the PK parameters described in Section 6.4.3 will be determined by a non-compartmental approach. For the purpose of the non-compartmental PK analysis, all plasma concentration BLQ values occurring prior to the first measurable concentration

will be set to zero. BLQ values occurring after the first measurable plasma concentration will be set to missing.

PK parameters will be summarized using descriptive statistics that includes the CV. Summaries of C_{max}, AUC_{0-24}, and $T_{1/2}$ will also include point estimates for the geometric mean and geometric CV. The geometric mean is calculated by computing the mean of the log-transformed concentration values and then presented in the original scale by calculating the anti-log of the mean result. The geometric CV is calculated as $100*sqrt[exp(\sigma^2)-1]$, where σ^2 is the variance of the log-transformed data. Listings of calculated PK parameters will be provided and will include the number of points used to calculate λ_z.

7.5 Safety Evaluation

Safety analysis for data collected during the RW phase will be based on the RW-FAS, to include all patients who receive at least one dose of double-blind study drug during the RW phase. Patients who do not complete the study, for whatever reason, will have all available data up until the time of termination included in the analysis. For safety analysis presented by study visit, the RW Baseline value will be defined as the last value reported prior to the first dose of study drug during the RW phase.

All safety summaries for data collected during the SAOL phase will be based on the SAOL-FAS, to include all patients who receive at least one dose of TA during the SAOL phase. For safety analysis presented by study visit, the SAOL Baseline value will be defined as the last value reported prior to the first dose of TA during the SAOL phase.

7.5.1 Extent of Exposure

Extent of exposure to study treatment during the RW phase will be summarized for the RW-FAS by treatment group and during the SAOL phase for the SAOL-FAS for all patients combined. The duration of exposure will be presented in days and calculated as the date of last dose of study drug minus the date of first dose of study drug, plus one. The date of last dose of study drug is the date reported on the End of Treatment form, for both the end of the SAOL phase and the end of treatment during the RW phase. Duration of exposure and total number of doses received will be summarized using descriptive statistics.

7.5.2 Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as those adverse events (AEs) with onset after the first dose of study drug or existing events that worsened after the first dose during the study. Treatment-emergent AEs occurring during the RW phase will be summarized by treatment group; TEAEs occurring during the SAOL phase will be summarized for all patients combined.

Treatment-emergent AEs occurring during the SAOL phase will be restricted to those events with onset on or after the start of dosing during the SAOL phase and prior to the start of dosing during the RW phase or 28 days after the last dose of study drug during

the SAOL phase, whichever occurs first. Events reported with a partial onset date (eg, month and year are reported but the day is missing) will be considered to be treatment-emergent if it cannot be confirmed that the event onset was prior to the first dose of study drug based on the available date entries.

Treatment-emergent AEs occurring during the RW phase will be restricted to those events with onset on or after the start of dosing during the RW phase and prior to the first dose of study drug during the open-label extension study or 28 days after the last dose of study drug during the RW phase, whichever occurs first.

Verbatim terms on case report forms will be mapped to preferred terms and system organ classes using the Medical Dictionary for Regulatory Activities (MedDRA, version 20.0).

Summaries that are displayed by system organ class and preferred terms will be ordered by descending incidence of system organ class and preferred term within each system organ class. Summaries displayed by preferred term only will be ordered by descending incidence of preferred term. Terms with the same incidence will be ordered alphabetically. Summaries of the following types will be presented, separately for the SAOL phase and RW phase:

- Overall summary of number of unique TEAEs and treatment-emergent serious adverse events (SAEs) and patient incidence of TEAEs meeting various criteria
- Patient incidence of TEAEs by MedDRA system organ class and preferred term
- Patient incidence of TEAEs by severity grade, MedDRA system organ class, and preferred term
- Patient incidence of TEAEs by relationship to study drug, MedDRA system organ class, and preferred term
- Patient incidence of TEAEs with severity grade ≥ 3 and related to study drug by MedDRA system organ class and preferred term
- Patient incidence of treatment-emergent adverse events of special interest (AESI) by MedDRA system organ class and preferred term
- Patient incidence of SAEs by MedDRA system organ class and preferred term

At each level of summarization (eg, any AE, system organ class, and preferred term), patients experiencing more than one TEAE will be counted only once. In the summary of TEAEs by severity grade, patients will be counted once at the highest severity reported at each level of summarization; in the summary of TEAEs by relationship, patients will be counted once at the closest relationship to study drug.

In addition to the TEAE incidence summaries described above, TEAE incidence will be corrected for differences in study drug exposure by using person-time in the denominator

to calculate incidence rates. Each patient's person exposure years (PEY) will be calculated by the following:

- As the date of onset of the TEAE of interest minus the date of first dose of study drug plus one, divided by 365.25, for those subjects who experience the TEAE of interest; or

- As the date of last dose of study drug plus 28 days, minus the date of first dose of study drug plus one, divided by 365.25, for those subjects who do not experience the TEAE of interest.

One PEY is the equivalent of one patient exposed to study drug for one year. Two patients who are exposed to study drug for half a year contribute one PEY. The total PEY of a treatment group is the sum of the person exposure years of each patient in that treatment group. Exposure-adjusted TEAE incidence will be summarized per 100 PEY as the number of patients with an event for whom person-time is available, divided by the total PEY for each treatment group and multiplied by 100. Exposure-adjusted TEAE incidence will be summarized separately for the SAOL phase and the RW phase.

Adverse event data will be presented in data listings by patient, treatment group, and event. Serious AEs, AEs leading to discontinuation of the study drug, and AESIs will be presented in separate data listings.

7.5.3 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events

All deaths during the study as reported on the End of Treatment (SAOL), End of Treatment (RW), Study Completion/Study Discontinuation, and Adverse Event eCRFs will be listed by patient, to include the primary cause of death. Serious AEs, AEs leading to study drug withdrawal, interruption, or dose reduction of the study drug, and AESIs will be provided in separate patient data listings. AESIs include events that constitute an injection site reaction, including pruritus, erythema, and swelling.

7.5.4 Clinical Laboratory Evaluation

All descriptive summaries of laboratory results will be based on data analyzed by the central laboratory and presented in Système International (SI) units, as suggested by the Center for Biologics Evaluation and Research and the Center for Drug Evaluation and Research *Position on Use of SI Units for Lab Tests* (Oct 2013). All data will be included in by-patient data listings. Laboratory measurements identified as abnormal (ie, outside the normal range) will also be listed separately by patient, laboratory test, and unit. In addition, normal ranges provided by the central laboratory will be presented in a separate listing.

Clinical laboratory measurements including serum chemistry, hematology, coagulation, and erythrocyte sedimentation rate will be summarized separately for the SAOL phase and the RW phase by treatment group. Descriptive statistics will be presented for

observed values and changes from each respective baseline at each visit where parameters were scheduled to be collected per the clinical study protocol.

Where applicable, laboratory results will be classified as "low," "normal," or "high" with respect to the parameter-specific reference ranges (ie, below the lower limit of the normal range, within the normal range, or above the upper limit of the normal range). Three-by-three contingency tables will be presented for each study phase and laboratory parameter to summarize the shift from the baseline category to the worst post-baseline measurement, defined as the value numerically farthest outside of the normal range across all post-baseline visits through the end of the study. Summary results will include the count and percentage of patients within each shift category and treatment group. Percentages will be based on the number of patients with a non-missing baseline value and at least one non-missing post-baseline value.

Where applicable, hematology, chemistry, and coagulation results for selected parameters will be assigned a toxicity grade based on the U.S. Department of Health and Human Services *Common Terminology Criteria for Adverse Events (CTCAE)*, version 5.0 (27 Nov 2017). If the quantitative criteria for grading are equivalent for two grades and the differentiation is described by clinical interventions, the clinical intervention component will not be considered and the highest CTCAE grade will be assigned. Similarly, death related to AE (ie, Grade 5) cannot be determined with available laboratory-based data collection and, thus, will not be summarized as a category. Laboratory parameters that include multiple sets of criteria for each direction (eg, separate criteria for potassium measures to assess hyperkalemia and hypokalemia) will be summarized separately to reflect each set of criteria.

Five-by-five contingency tables will be presented by study phase for lab tests where toxicity grading can be applied, to summarize the shift from the baseline grade to the worst post-baseline grade. Grades will be presented as none (Grade 0; ie, measurement did not meet any CTCAE criteria for Grades 1 through 4), mild (Grade 1), moderate (Grade 2), severe (Grade 3), or life-threatening (Grade 4). Summary results will include the count and percentage of patients within each shift category. Percentages will be based on the number of patients with a non-missing baseline value and at least one non-missing post-baseline value.

7.5.5 Vital Signs, Physical Findings, and Other Observations Related to Safety

7.5.5.1 Vital Signs

Vital sign parameter measurements will be summarized by study phase and treatment group. Descriptive statistics will be presented for results and change from phase-specific baseline at each visit where parameters were scheduled to be collected. Daily temperatures will be listed. In addition, the number of days where fever is present, defined as temperature above 38 °C, will be presented by month for each subject.

7.5.5.2 Local Tolerability Index Assessment

The local tolerability index assesses severity of symptoms for each dosing day. For each symptom, the worst severity reported for any day will be summarized with counts and percentages of patients in each category by study phase and treatment group.

7.5.5.3 Skin Rash

The presence and severity of a skin rash will be evaluated at baseline and each post-baseline visit. The number and percent of patients will be presented by study phase, visit, and treatment group for patients with a skin rash present and for each level of severity (mild, moderate, or severe). At post-baseline visits within each phase, the number and percent of patients with the rash at the phase-specific baseline still present and by change category will be presented by treatment group.

7.5.5.4 Immunogenicity Evaluation

Blood levels of antibodies to IL-18BP antibodies TA (anti-rh-IL-18BP) will be summarized by treatment group, separately for each study phase. Descriptive statistics will be presented for results and change from phase-specific baseline at each visit where blood samples were scheduled for collection. Blood levels for anti-rh-IL-18BP will also be presented graphically over time for each individual patient.

7.5.5.5 Physical Examination

Results of the physical examination will be presented in patient data listings by patient, study visit, and body system.

7.5.5.6 Ophthalmologic Examination

Results of the ophthalmologic examination, to include a uveitis assessment and other anterior chamber abnormal pathology findings, will be presented in patient data listings by patient and study visit.

7.5.5.7 Concomitant and Rescue Medications

Medications will be coded using the WHODDE (version March 1, 2017). Medications entered on the eCRF will be mapped to ATC drug class (level 4) and drug name.

Concomitant medications will include all medications reported on the Concomitant Medications eCRF. Medications will be mapped to the study phase received (SAOL/RDBPC or RW) based on the start and end dates. All medications will be listed, to include the derived study phase when the medication was received.

A concomitant medication during each phase is defined as any medication administered on or after the date of the first dose of study drug for that phase (SAOL/RDBPC or RW). A medication may be defined as being received during both the SAOL/RDBPC and RW phases. If it cannot be determined whether a medication was received after the start of

study drug dosing for each phase, it will be considered concomitant during that phase. Medications received during the study and identified as rescue immunosuppression on the eCRF will be summarized separately for the SAOL and RW phases, and listed.

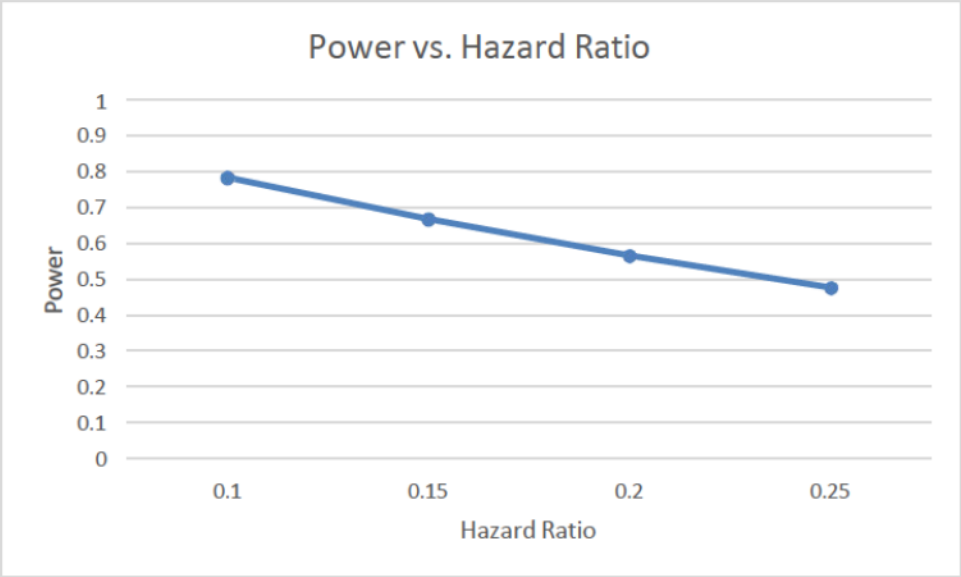
For concomitant and rescue medication use, the number and percentage of patients receiving any medication will be summarized by treatment group, as will the number and percentage receiving any medication by ATC drug class and generic drug name. Patients reporting use of more than one medication at each level of summarization (any medication received, ATC class, and generic drug name) will be counted only once. ATC class terms will be displayed by descending order of incidence, as will generic drug names within each ATC class. Terms with the same incidence will be ordered alphabetically. The summary of rescue medications received will also present the number and percentage of patients who respond to at least one rescue treatment.

7.6 Determination of Sample Size

The primary endpoint of the study is time to first occurrence of disease reactivation (including full and partial disease reactivation) during the RW phase of the study. There is limited historical data for this disease condition, for either a control or an experimental test product, however the frequent recurrence of disease reactivations was expected in the control arm (Wada 2014). Originally, an 8-week event rate is assumed to be 90% in the placebo group and a considerable reduction in disease reactivations for patients receiving TA was expected. However, a lower event rate was observed from the currently enrolled (still blinded) patients. Thus, the RW phase is prolonged to a maximum of 16 weeks to better assess the effect of TA withdrawal. The power calculations below will assume the same 90% event rate in the SOC arm at 16 weeks. Figure 1 below displays the power achieved for a two-arm study and assumes the following:

- Two-sided log-rank test, with input based on the proportion of patients surviving (ie, event-free) at 16 weeks
- Two-sided alpha = 0.05
- A 1:1 randomization of 10 patients (ie, 5 patients per arm)
- The proportion of patients surviving (ie, event-free) in the control arm at Week 16 of the RW phase is 0.1, which equates to a 90% event rate
- No patients in either group drop out of the study prior to experiencing a disease reactivation or completing the 16 week treatment period as planned
- Total study time is 32 weeks, including a 16-week accrual period and a 16 week follow-up time
- Varying levels of a hazard ratio, from 0.1 to 0.25 in 0.05 increments

Figure 1 Power vs. Hazard Ratio



A hazard ratio of 0.1 equates to an assumed 16-week event rate in the TA group of 20.6% (ie, a survival event-free rate of 79.4%). These assumptions achieve 78.4% power based on a two-sided log rank test with an overall sample size of 10 patients to be randomized in the RW phase (5 in the placebo group and 5 in the TA group). A hazard ratio of 0.1 results in an assumed number of total events among 10 patients to be 6.7, with 4.9 occurring in the placebo group and 1.8 occurring in the TA group.

No power calculations were performed for the SAOL phase of the study, as analysis will be descriptive in nature.

7.7 Changes in the Conduct of the Study or Planned Analyses

Changes to the planned analyses relative to what was outlined in the clinical study protocol (Version 6.0) include the following:

- A description for how to handle situations where clinical or laboratory signs of inflammation meeting the criteria of disease reactivation for the primary endpoint are present but are determined by the treating physician to be related to an infection rather than disease reactivation and are resolved after repeat assessments within the visit window, has been added to Section 7.4.2.2, but is not included in the clinical study protocol.

- A correction to the RW Per-Protocol Analysis Set (RW-PPAS) listed in the protocol is included below.

The study under the original protocol and through Amendment 2 of the protocol (Version 3.0) was designed with an 18-week RDBPC phase, followed by an 8-week RW phase for further efficacy and safety evaluation. Three patients were enrolled in the study under these protocol versions and each either completed the RDBPC phase and RW phase or

terminated the study early. As described in Section 7.1.6, patients randomized to the TA group during the RDBPC phase of the original study design will be included in analyses for the updated study design, to be combined with those patients enrolled under Amendment 4 (Version 4.0) and any subsequent amendments. Protocol version 5.0 and subsequent amendments extended the RW phase from 8 weeks with last the RW visit at Week 26 [V11] to 16 weeks, with the last RW visit at Week 34 [V13]. Accordingly, the RW Per-Protocol Analysis Set (RW-PPAS) is adapted to include all patients in the RW-FAS who complete the RW phase up through Week 26 (V11) for protocol version 4.0/4.1 or Week 34 (V13) for protocol versions 5.0 and 6.0 or have disease reactivation and for whom no relevant protocol deviations were documented.

Protocol version 5.0 and subsequent amendments identify response to therapy during the SAOL phase as a key secondary endpoint. However, since this key secondary endpoint is only being presented descriptively without any statistical testing, it is not included in the heirchical list for fixed-sequence testing described in Section 7.4.5.5.

Appendix A includes a description of the mapping of the mAIDAI components into the Objective mAIDAI total score.

8 REFERENCE LIST

Wada, T. et al. Sustained elevation of serum interleukin-18 and its association with hemophagocytic lymphohistiocytosis in XIAP deficiency. Cytokine 65, 74-78 (2014).

White, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrics*, 48, 817-838 (1980).

APPENDIX A

Objective mAIDAI total score mapping:

mAIDAI score			**Objective mAIDAI components for efficacy evaluation**	
	Score			**Score**
Fever > 100.4	☐ 0 (No) - ☐ 1 (Yes)	→	Fever > 100.4	☐ 1 (Yes)
Abdominal Pain / Colic	☐ 0 (No) - ☐ 1 (Yes)			
Nausea / Vomiting	☐ 0 (No) - ☐ 1 (Yes)			
Diarrhea	☐ 0 (No) - ☐ 1 (Yes)	→	Diarrhea	☐ 1 (Yes)
Transaminitis	☐ 0 (No) - ☐ 1 (Yes)	→	Transaminitis and/or hepatomegaly	☐ 1 (Yes)
Organomegaly If yes, specify ☐ Hepatomegaly ☐ Splenomegaly ☐ Adenopathy	☐ 0 (No) - ☐ 1 (Yes)	→	Splenomegaly and/or adenopathy	☐ 1 (Yes)
Rash	☐ 0 (No) - ☐ 1 (Yes)	→	Rash	☐ 1 (Yes)
Uveitis 3+ / 4+	☐ 0 (No) - ☐ 2 (Yes)	→	Uveitis	☐ 1 (Yes)
Uveitis 1+ / 2+	☐ 0 (No) - ☐ 1 (Yes)			
Arthralgia	☐ 0 (No) - ☐ 1 (Yes)			
Arthritis	☐ 0 (No) - ☐ 1 (Yes)	→	Arthritis	☐ 1 (Yes)
Radiologic CNS Involvement	☐ 0 (No) - ☐ 1 (Yes)	→	Radiologic CNS Involvement	☐ 1 (Yes)
CNS Symptoms	☐ 0 (No) - ☐ 1 (Yes)			
Cytopenia (HgB < 9.0, PLT < 100, WBC < 3.0)	☐ 0 (No) - ☐ 1 (Yes)	→	Cytopenia (HgB < 9.0, PLT < 100, WBC < 3.0)	☐ 1 (Yes)
mAIDAI Score at this visit (total sum of all scores from each category above):	_____		**Number of affected end organs = objective mAIDAI components:**	_____