Statistical Analysis Plan

Protocol Title:	A Phase 2 Long-Term Extension (LTE) Study to Evaluate The Safety and Efficacy of Efavaleukin Alfa in Subjects With Moderately to Severely Active Ulcerative Colitis
Short Protocol Title:	Long-term Safety and Efficacy of Efavaleukin Alfa in Subjects With Moderately to Severely Active Ulcerative Colitis
Protocol Number:	20210210
NCT Number:	NCT05672199
Authors:	[redacted]
Sponsor:	Amgen Inc. One Amgen Center Drive, Thousand Oaks, CA, 91320, USA
SAP Date:	**Document Version** — **Date** **Original (v1.0)** — 22 August 2023 **Amendment 1 (v2.0)** — **19 February 2025**

Version Number	Date (DDMMMYYYY)	Summary of Changes, including rationale for changes
Original (v1.0)	22AUG2023	Not Applicable
Amendment 1 (v2.0)	**19FEB2025**	**Updates to be consistent with protocol amendment 2 and decision of study 20170104 IA3 to terminate the study:** • **Sample size reduced to 92 from 140** • **Exploratory endpoints removed**

Table of Contents

Table of Contents 3

1. Introduction 6

2. Objectives, Endpoints/Estimands and Hypotheses 6
   - 2.1 Objectives and Endpoints/Estimands 6
   - 2.2 Hypotheses and/or Estimations 7

3. Study Overview 7
   - 3.1 Study Design 7
   - 3.2 Sample Size 7

4. Covariates and Subgroups 7
   - 4.1 Planned Covariates 7
   - 4.2 Subgroups 8

5. Definitions 8
   - 5.1 Basic Definition 8
   - 5.2 Study Points of Reference 10
   - 5.3 Arithmetic Calculations 10
   - 5.4 Definition of Study Endpoints 11

6. Analysis Sets 13
   - 6.1 Full Analysis Set 13
   - 6.2 Safety Analysis Set 13
   - 6.3 Anti-drug Antibody Analysis Set 13

7. Planned Analyses 13
   - 7.1 Final Analysis 13

8. Data Screening and Acceptance 14
   - 8.1 General Principles 14
   - 8.2 Data Handling and Electronic Transfer of Data 14
   - 8.3 Handling of Missing and Incomplete Data 14
   - 8.4 Detection of Bias 15
   - 8.5 Outliers 15
   - 8.6 Distributional Characteristics 15
   - 8.7 Validation of Statistical Analyses 15

9. Statistical Methods of Analysis 16
   - 9.1 General Considerations 16
   - 9.2 Subject Accountability 16
   - 9.3 Important Protocol Deviations 16
   - 9.4 Demographic and Baseline Characteristics 16
   - 9.5 Efficacy Analyses 18
     - 9.5.1 Analyses of Primary Efficacy Endpoint(s)/Estimand(s) 18

9.5.2 Analyses of Secondary Efficacy Endpoint(s)/Estimand(s) 18
9.6 Safety Analyses 19
9.6.1 Adverse Events 19
9.6.2 Laboratory Test Results 19
9.6.3 Vital Signs 20
9.6.4 Physical Measurements 20
9.6.5 Electrocardiogram 20
9.6.6 Antibody Formation 20
9.6.7 Exposure to Investigational Product 20
9.6.8 Exposure to Other Protocol-permitted Therapy 20
10. Changes From Protocol-specified Analyses 21
11. Literature Citations / References 21
12. Appendices 22
Appendix A. Reference Values/Toxicity Grades 22
Appendix B. Analytical Windows 23
Appendix C. Handling of Dates, Incomplete Dates and Missing Dates 27
Appendix D. Mayo Score Categories 28

List of Abbreviations

Abbreviation	Explanation
ASA	aminosalicylates
CTCAE	Common Terminology Criteria for Adverse Events
DMC	data monitoring committee
ECG	electrocardiogram
eCRF	electronic case report form
EOS	End of Study
EOIP	End of Investigational Product
ET	Early Termination
FAS	Full Analysis Set
FDA	Food and Drug Administration
IL-2	interleukin-2
IV	Intravenous
JAK	Janus kinase
LTE	Long-term extension
NCT	National Clinical Trials
NK	natural killer
[redacted]	
S1P modulator	Sphingosine-1-phosphate receptor modulator
SAP	Statistical analysis plan
SC	Subcutaneous
TEAE	Treatment Emergent Adverse Event
TNF	tumor necrosis factor
UC	ulcerative colitis

1. Introduction

The purpose of this Statistical Analysis Plan (SAP) is to provide details of the statistical analyses that have been outlined within the protocol amendment **2** for study 20210210, Efavaleukin Alfa (AMG 592) dated **19 July 2024**. The scope of this plan includes the final analysis that is planned and will be executed by the Amgen Global Biostatistical Science department unless otherwise specified.

2. Objectives, Endpoints/Estimands and Hypotheses

2.1 Objectives and Endpoints/Estimands

Objectives	Endpoints
Primary	
• To evaluate the long-term safety and tolerability of efavaleukin alfa in subjects with moderate to severe ulcerative colitis (UC)	• Treatment-emergent adverse events
Secondary	
• To evaluate the effect of efavaleukin alfa long-term treatment on clinical response	• Clinical response at week 52 and 104
• To evaluate the effect of efavaleukin alfa long-term treatment on clinical remission	• Clinical remission at week 52 and 104
• To evaluate the effect of efavaleukin alfa long-term treatment on durable clinical remission	• Durable clinical remission at week 52 and 104
• To evaluate the effect of efavaleukin alfa long-term treatment on endoscopic remission	• Endoscopic remission at week 52 and 104
• To evaluate the effect of efavaleukin alfa long-term treatment on histologic remission	• Histologic remission at week 52 and 104
• To evaluate the effect of efavaleukin alfa long-term treatment on corticosteroid-free remission	• Corticosteroid-free remission at week 52 and 104 in subjects receiving corticosteroids at randomization of the phase 2 parent Study 20170104
• To evaluate the effect of efavaleukin alfa long-term treatment on combined endoscopic and histologic remission	• Combined endoscopic and histologic remission at week 52 and 104
• To evaluate the effect of efavaleukin alfa long-term treatment on symptomatic remission	• Symptomatic remission at week 52 and 104

• To evaluate the effect of efavaleukin alfa long-term treatment on change in histological score	• Change from baseline of Study 20170104 in histological score (Geboes) at week 52 and 104

2.2 Hypotheses and/or Estimations

No formal statistical hypothesis will be tested in this study, all analyses will be descriptive.

3. Study Overview

3.1 Study Design

The study is a phase 2 Long Term Extension (LTE) study for subjects who completed the 52 week phase 2 parent dose-finding study (Study 20170104). Subjects will enroll into the LTE (Study 20210210) in a blinded manner on the same dose they were receiving at the completion of their participation in Study 20170104. This study will evaluate the long-term safety and clinical efficacy of efavaleukin alfa in subjects entering Study 20210210 with moderately to severely active UC.

Subjects will continue to receive the dose of investigational product (placebo or efavaleukin alfa) that they were receiving at the completion of their participation in Study 20170104:

- [redacted] µg SC [redacted]
- [redacted] µg SC [redacted]
- [redacted] µg SC [redacted]
- Placebo SC [redacted] (if subject was a Study 20170104 [redacted] placebo responder)

An independent Data Monitoring Committee (DMC) will oversee safety and overall conduct of the study. The DMC will review LTE (Study 20210210) study data per the schedule as outlined in the DMC charter.

3.2 Sample Size

Approximately 140 subjects **were** expected to be enrolled in the LTE study. **The sample size is reduced to 92 subjects as a result of termination of the parent study due to futility.**

4. Covariates and Subgroups

4.1 Planned Covariates

No covariates will be considered in this study because no modeling will be conducted.

4.2 Subgroups

No subgroup analysis is planned for the final analysis of the study.

5. Definitions

5.1 Basic Definition

Investigational Product (IP):

The term investigational product is used in reference to efavaleukin alfa and placebo.

Enrollment Date:

The enrollment date is the date on which a subject received his/her first dose of IP in the Study 20210210.

First Dose Date:

The first dose date is the date of administration of the first dose of IP in the Study 20210210.

Duration of IP Exposure:

The duration of IP exposure will be derived as the date of last IP administration plus 13 days (14 day window - 1), end of study date, whichever occurs first, minus the date of first IP administration + 1 day.

Completion of IP:

A subject will be considered as completing IP if End of IP reason is “Completed” on the End of IP eCRF.

Actual Treatment Received:

The actual treatment received is the IP the subject actually received during Study 20210210 through the end of study. In cases where a subject received both efavaleukin alfa and placebo, the actual treatment received will be the efavaleukin alfa dose. In cases where a subject received multiple efavaleukin alfa doses, the actual treatment received will be the maximum efavaleukin alfa dose received. **In cases where a subject received an efavaleukin alfa dose higher than the highest dose specified in the protocol, the actual treatment received will be the highest dose.**

Treatment-emergent adverse event (TEAE):

Treatment-emergent adverse events are events categorized as Adverse Events (AEs) starting on or after first dose of IP in Study 20210210 as determined by “Did event start

before first dose of investigational product?" equal to "No" or missing on the Events eCRF and up to the end of study date.

Treatment-related Adverse Event (TRAE):

A treatment-related AE is any TEAE in Study 20210210 with the relationship flag on the Events eCRF indicating there is a reasonable possibility that the event may have been caused by investigational medicinal product as determined by the investigator. In the unlikely event that the relationship is missing, the treatment-emergent event will be considered treatment-related and documented in a footnote of the treatment-related event summary table**.**

Exposure-Adjusted Adverse Events:

The exposure for the analysis of exposure-adjusted TEAEs and serious TEAEs is defined for each subject as **(EOS – first dose date +1)**.

Baseline medication:

Baseline medication is defined as any medication with start date on or before study day 1 and ongoing while on study.

Concomitant medication:

Concomitant medication is defined as any medication with start date prior to study day 1 and ongoing while on study or any medication with start date on or after study day 1 and up to the end of study.

Prior medication:

Prior medication is defined as any medication with an end date prior to study day 1 of the Study 20170104.

Biologic Therapy Classes:

Biologic therapy classes include:

- Anti-TNF agents: **Infliximab, Adalimumab, Golimumab**
- Integrin receptor antagonists: **Vedolizumab**
- **Interleukin inhibitors: Ustekinumab (an anti-interleukin-12/23 antibody), Mirikizumab (a monoclonal antibody that targets the p19 subunit of interleukin-23), Risankizumab (a monoclonal antibody directed against the p19 subunit of IL-23)**

Targeted Small Molecules:

Targeted small molecules include:

- **Janus kinase (JAK) inhibitors: Tofacitinib and Upadacitinib**
- **Sphingosine-1-phosphate (S1P) receptor modulators: Ozanimod and Etrasimod**

5.2 Study Points of Reference

Baseline:

Baseline fecal calprotectin is the last measurement taken on or before the date of the first dose of IP in Study 20170104.

Baseline for all other parameters is the last measurement from Study 20170104 for the endpoint of interest taken before the first dose of IP in Study 20170104. If the measurement is taken on the same day as the first dose and the exact measurement time relative to the first dose is unknown, it will be assumed to have been taken prior to the first dose of IP.

Study Day 1:

Study day 1 for each subject is the date of first IP administration in Study 20210210.

Study Day:

Study day for each subject is defined as (day of interest - study day 1) + 1 for dates on or after study day 1, or (day of interest - study day 1) for dates prior to study day 1.

5.3 Arithmetic Calculations

Change from Baseline:

The arithmetic difference between a post-baseline value and baseline for a given timepoint:

(post-baseline value – baseline value).

Percent Change from Baseline:

The change from baseline divided by baseline value and multiplied by 100:

(change from baseline / baseline) * 100.

If baseline is zero, the percent change from baseline will be set to zero.

Fold Change from Baseline:

Fold change from baseline equals the post-baseline value divided by the baseline value**.** If the change from baseline is not equal to 0 and the baseline value is 0 then fold change

is not defined. If the change from baseline is equal to 0 and the baseline value is also 0 then fold change is 1.

5.4 Definition of Study Endpoints

Individual Mayo Subscore:

There are 4 Individual Mayo subscores; stool frequency, rectal bleeding, locally read endoscopy findings, and physician's global assessment (PGA), with each subscore graded semi-quantitatively on a score of 0 to 3.

Further explanation given in Appendix **D** for the full definition of mayo subscores and calculation.

Total Mayo Score:

The total Mayo score is a composite index of the sum of the individual Mayo subscores, with each item graded semi-quantitatively on a score of 0 to 3, for a maximal total score of 12. The total Mayo score ranges from 0 to 12 points.

Modified Mayo Score:

The modified Mayo score is the total Mayo score without the physician's global assessment subscore and ranges from 0 to 9 points.

Partial Mayo Score:

The partial Mayo score is the total Mayo score without the endoscopy subscore and ranges from 0 to 9 points.

Clinical Remission:

Clinical remission is defined as modified Mayo score 0 to 2 including rectal bleeding subscore of 0, a stool frequency subscore of 0 or 1, and a locally read endoscopy subscore of 0 or 1 (modified so that 1 does not include friability).

Clinical Response:

Clinical response is defined as a decrease from baseline in the modified Mayo score of $\geq$ 2 points and at least 30% reduction from baseline, and a decrease in the rectal bleeding subscore of $\geq$ 1 or an absolute rectal bleeding subscore of 0 or 1.

Durable Clinical Remission:

Durable clinical remission is defined as **the maintenance of** clinical remission at **both** week 52 **(2 years after initial IP)** and 104 **(3 years after initial IP)** time points among subjects in clinical remission at [redacted] of Study 20170104.

Endoscopic Remission:

Endoscopic remission is defined as Mayo locally read endoscopy subscore of 0 or 1 (modified so that a score of 1 does not include friability).

Symptomatic Remission:

Symptomatic remission is defined as Mayo stool frequency subscore of 0 or 1 and rectal bleeding subscore of 0.

Combined endoscopic and histologic remission:

Combined endoscopic and histologic remission is defined as combined endoscopic remission (Mayo locally read endoscopy subscore of 0 or 1) and histologic remission of the colon tissue (Geboes Score < 2.0; no neutrophils in the epithelium crypts or lamina propria and no increase in eosinophils, no crypt destruction and no erosions, ulcerations or granulation tissue [Geboes, 2000]).

Corticosteroid-free Remission:

For subjects who received corticosteroids at randomization of Study 20170104, corticosteroid-free remission is defined as achieving clinical remission at week 52 **(2 years after initial IP)** and week 104 **(3 years after initial IP)** and successfully tapering corticosteroids to 0mg per day for the entire 12 weeks prior to the date of the week 52 **(2 years after initial IP)** and week 104 **(3 years after initial IP)** endoscopy respectively.

Rectal Bleeding and Stool Frequency Subscores:

Rectal bleeding and number of stools per day will be recorded in the Mayo Daily Symptom Diary by the patient. The stool frequency subscore will be calculated based on the number of stools per day more than the reference number of stools. The reference number of stools is defined as the number of stools in a 24-hour period when in remission from UC prior to enrolment into the Study 20170104. If the subject has never achieved remission, the number of stools per day before initial onset of signs and symptoms of UC will be used.

The subscores will be derived from the most recent 3 days of data entered within the last 7 days prior to the visit, or within 7 days prior to initiation of bowel preparations for visits which include an endoscopy. The visit will be defined as the date of the PGA assessment. If PGA is missing, the target assessment date for that visit per Appendix B will be used.

Rules for consideration of visit dates are defined in Appendix E.

Worst Case Imputation:

Worst case imputation will be used for missing postbaseline values for Geboes score, total Mayo score, partial Mayo score, and individual Mayo subscores. Missing postbaseline values at any visit will be imputed using the highest score observed across baseline and all postbaseline visits.

6. Analysis Sets

6.1 Full Analysis Set

The full analysis set includes all enrolled subjects. Data will be analyzed based on planned treatment group. **Subject data from site 66164 are not included in the analysis set because the site closed due to noncompliance with the protocol and FDA regulations.**

6.2 Safety Analysis Set

The safety analysis set includes all enrolled subjects who received at least 1 dose of investigational product. Data will be analyzed based on actual treatment group. **Subject data from site 66164 are not included in the analysis set because the site closed due to noncompliance with the protocol and FDA regulations.**

6.3 Anti-drug Antibody Analysis Set

Anti-drug Antibody Analysis Set is defined as the subset of subjects in the Safety Analysis Set who had at least 1 evaluable anti-drug antibody test. Immunogenicity data will be analyzed according to the actual treatment received. **Subject data from site 66164 are not included in the analysis set because the site closed due to noncompliance with the protocol and FDA regulations.**

7. Planned Analyses

7.1 Final Analysis

Final analysis activities will commence upon achieving the end of study. The final analysis will include analysis of efficacy and safety data.

Data will be subject to ongoing checks for integrity, completeness, and accuracy in accordance with the Data Management Plan with the expectation that all outstanding data issues are resolved ahead of the final lock.

8. Data Screening and Acceptance

8.1 General Principles

The objective of the data screening is to assess the quantity, quality, and statistical characteristics of the data relative to the requirements of the planned analyses.

8.2 Data Handling and Electronic Transfer of Data

Global Study Operations-Data Management (GSO-DM) department will provide all data to be used in the planned analyses. This study will use the RAVE database. All laboratory values from the central laboratory will be transferred to GSO-DM. All PRO data, endoscopy results (including Mayo endoscopic subscore and Geboes score from the mucosal biopsy), and individual Mayo subscores will be transferred to GSO-DM. All other data will be captured on the eCRF.

8.3 Handling of Missing and Incomplete Data

Subjects may have missing specific data points for a variety of reasons. In general, data may be missing due to a subject's early withdrawal from study, a missed visit, or non-evaluability of a specific clinical measurement at its planned clinical visit. Missing data will not be imputed, except for the data points described below.

Non-responder imputation will be used to impute missing data for the following binary endpoints: clinical remission, clinical response, endoscopic remission, symptomatic remission, combined endoscopic and histologic remission, corticosteroid-free clinical remission.

If a subject has missing data for any of these endpoints, (including missing one or more subscores contributing to the endpoint) at any scheduled visit, the subject will be considered as not meeting the endpoint for that specific visit.

Missing data for the following continuous endpoints will be imputed with worst case imputation: change in Geboes score, total Mayo score, partial Mayo score, and individual Mayo subscores.

Observed subscores will be used in the calculation of total and partial mayo score where available. If one or more individual subscores is missing at a given post-baseline timepoint, the missing subscore will be imputed using the worst (highest) observed score across all available baseline and post-baseline visits up

to [redacted] The total/partial mayo score will be calculated using the imputed subscores, along with observed subscores, if available for that visit.

Laboratory measurements that are below the lower quantification limits will be considered equal to the lower limit of quantification for all analyses unless explicitly noted otherwise.

Incomplete UC diagnosis date will be imputed as follows:

- If day is missing but month and year are present, impute to the last day of the month
- If day and month are missing and year is present, impute to the 31st December of that year

Missing and incomplete dates of medical history, procedures, AE, concomitant medication and hospitalization dates will be imputed as outlined in Appendix C.

8.4 Detection of Bias

Important protocol deviations and early withdrawal from treatment and from study may bias the results of the study. The incidence of these factors will be assessed and reason for early withdrawals will be tabulated.

8.5 Outliers

Scatter plots will be examined to identify potential outliers in any of the continuous variables and frequencies of the categorical data will be examined to identify questionable values. The validity of any questionable values will be verified, and observations found to be due to data entry errors will be queried. Potential outliers that are not due to data entry error will be included in the analysis.

8.6 Distributional Characteristics

Not applicable for this study.

8.7 Validation of Statistical Analyses

Programs will be developed and maintained and output will be verified in accordance with current risk-based quality control procedures.

Tables, figures, and listings will be produced with validated standard macro programs where standard macros can produce the specified outputs.

The production environment for statistical analyses consists of Amgen-supported versions of statistical analysis software; for example, the SAS System version 9.4 or later.

9. Statistical Methods of Analysis

9.1 General Considerations

Statistical analyses in this Study 20210210 are descriptive in nature. Summary statistics for continuous variables will include the number of subjects, mean, median, **first quartile (Q1), third quartile (Q3)**, standard deviation, minimum, and maximum. For categorical variables, the frequency and percentage will be given.

9.2 Subject Accountability

The number and percent of subjects who were enrolled, received investigational product, completed investigational product, discontinued investigational product (including reasons for discontinuing), completed study, discontinued study (including the reason for discontinuing) will be summarized overall and by planned treatment group.

Key study dates for the first subject enrolled, last subject enrolled, last subject's end of investigational product, and last subject's end of study will be presented.

9.3 Important Protocol Deviations

Protocol Deviations (IPDs) categories are defined by the study team before the first subject’s initial visit and updated during the IPD reviews throughout the study prior to final database lock. These definitions of IPD categories, subcategory codes, and descriptions will be used during the course of the study. Eligibility deviations are defined in the protocol.

The final IPD list will be used to produce the summary of IPDs table and the list of subjects with IPDs.

9.4 Demographic and Baseline Characteristics

The following demographic, baseline characteristics, baseline disease characteristics and baseline and prior medications of interest will be summarized descriptively by planned treatment group.

All demographics and baseline variables except baseline medication will be sourced from the Study 20170104.

Demographics:

- Age (years)

- Age groups (≥ 18 to < 50, ≥ 50 and 18 – 64, 65 – 74,75 – 84)
- Sex (male, female)
- Race (American Indian or Alaska, Native Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Geographic region (Eastern Europe, Asia Pacific, Rest of World)

Baseline Characteristics:

- Height (cm)
- Weight (kg)
- Body Mass Index (BMI) (kg/m^2)

Baseline Disease Characteristics:

- Modified Mayo score
- Total Mayo score
- Individual Mayo subscores
- Partial Mayo score
- **Disease localization (proctosigmoiditis, left-sided colitis, extensive colitis, pancolitis)**
- **Duration of UC (years from date of diagnosis to first dose date)**
- **Fecal calprotectin**

Baseline medications:

- Oral 5-aminosalicylates (5-ASA)
- Corticosteroid
- Immunomodulator (azathioprine, 6-mercaptopurine, methotrexate)

Prior medication history:

Prior experience, including reason for stopping, for the following medication classes:

- conventional therapies (azathioprine, 6-mercaptopurine, methotrexate, corticosteroids)
- biologic therapy (including number of prior biologic therapy classes and total number of prior biologics)
- anti-TNF therapy
- targeted small molecule therapy (ie. Janus kinase [JAK] inhibitor and S1P modulator)

9.5 Efficacy Analyses

The efficacy analyses are summarized in Table 9-1and analyses of efficacy endpoints are described from Section 9.5.1 to Section 9.5.3.

Table 9-1. Secondary Efficacy Endpoint Summary Table

Endpoint	Primary Summary and Analysis Method - FAS
• Clinical response at week 52 and 104 • Clinical remission at week 52 and 104 • Durable clinical remission at week 52 and 104 • Endoscopic remission at week 52 and 104 • Histologic remission at week 52 and 104 • Corticosteroid-free remission at week 52 and 104 in subjects receiving corticosteroids at randomization of the phase 2 parent Study 20170104 • Combined endoscopic and histologic remission at week 52 and 104 • Symptomatic remission at week 52 and 104 • Change from baseline of Study 20170104 in histological score (Geboes) at week 52 and 104	Summary statistics

9.5.1 Analyses of Primary Efficacy Endpoint(s)/Estimand(s)

Not applicable for this study.

9.5.2 Analyses of Secondary Efficacy Endpoint(s)/Estimand(s)

The analysis of secondary efficacy endpoints is based on the Full Analysis Set (FAS) described in Section 6.1 and the subjects with missing data will be imputed using nonresponder imputation.

Summary statistics will be provided by planned treatment group.

9.6 Safety Analyses

9.6.1 Adverse Events

The study will include events starting on or after first dose of investigational product in the Study 20210210. The adverse events will be summarized by actual treatment group.

The Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or later will be used to code all events categorized as adverse events to a system organ class and a preferred term. The severity of each adverse event will be graded using CTCAE criteria version 5.0 or later (Appendix A).

Subject incidence of all TEAEs, serious TEAEs, TEAEs leading to discontinuation from investigational product or other protocol-required therapies, fatal TEAEs, treatment-related TEAEs, treatment-related serious TEAEs, treatment-related TEAEs leading to discontinuation from investigational product, treatment-related fatal TEAEs and TEAEs of interest will be tabulated by system organ class and preferred term.

Exposure-adjusted subject incidence of TEAEs and serious TEAEs will be summarized.

Subject incidence of events of interest (standardized MedDRA queries and/or Amgen Medical Queries) will also be summarized according to their categories and preferred term. Events of interest could include but are not limited to Hypersensitivity, Tachyarrhythmias, Tachypnea, Hematopoietic Cytopenia, Drug Reaction with Eosinophilia, Leucocyte Changes, Infection and Infestation, Cytokine Release Syndrome, and Injection Site Reaction. Number of episodes and duration of treatment-emergent injection site reactions will be further summarized by treatment group.

Overall summary of subject incidence of TEAEs by preferred term and worst severity grade will be provided by anti-AMG592 and Anti-IL-2 antibody status.

9.6.2 Laboratory Test Results

Change and percent change from baseline in select clinical laboratory test results will be summarized over time by each treatment group. In addition, shift tables, from baseline to the worst on-study laboratory toxicity based on the CTCAE version 5 will be presented.

Subject incidence of worst post-baseline of eosinophils counts and percentages by the laboratory normal range will be tabulated by treatment group.

9.6.3 Vital Signs

The actual value, change and percent change from baseline in vital signs will be summarized over time by treatment arm for systolic blood pressure, diastolic blood pressure, heart rate, respiratory rate, and temperature.

9.6.4 Physical Measurements

The actual value, change and percentage change from baseline in weight and BMI will be summarized by treatment arm and by scheduled visit.

9.6.5 Electrocardiogram

The ECG measurements from this clinical study will be performed as per standard of care for routine safety monitoring, rather than for purposes of assessment of potential QTc effect. Since these evaluations may not necessarily be performed under the rigorous conditions expected to lead to meaningful evaluation of QTc data; summaries and statistical analyses of ECG measurements are not planned, and these data would not be expected to be useful for meta-analysis with data from other trials.

9.6.6 Antibody Formation

The incidence and percentage of subjects who develop anti-efavaleukin alfa antibodies and antibodies to IL-2 (binding and if positive, neutralizing) at any time will be tabulated by actual treatment received. Furthermore, the incidence and percentage of subjects with treatment-boosted anti- efavaleukin alfa antibodies will also be summarized.

9.6.7 Exposure to Investigational Product

Descriptive statistics will be produced to describe the exposure to IP by treatment arm. Summary statistics will be provided for the total number of doses administered, total dose received, and total duration of IP exposure by treatment arm.

9.6.8 Exposure to Other Protocol-permitted Therapy

Protocol-permitted therapy includes oral 5-ASA compounds, immunomodulators, and oral corticosteroids.

Descriptive statistics of number and percentage of subjects with dose modifications, including steroid tapering of corticosteroids will be produced by treatment arm.

The number and percentage of subjects taking oral 5-ASA compounds, immunomodulators, and oral corticosteroids will be provided.

10. Changes From Protocol-specified Analyses

- **The third interim analysis conducted for Study 20170104 in August 2024 revealed that the study met futility criteria, leading to its termination as well as the termination of this LTE study. As a result of this decision, exploratory endpoints originally included in the Statistical Analysis Plan (SAP) are being removed. These endpoints will not be discarded entirely; instead, they will be documented in a Supplemental Statistical Analysis Plan (SSAP) and will not be documented in CSR.**
- **Subject data from site 66164 are not included in the analysis set because the site closed due to noncompliance with the protocol and FDA regulations. The following listing for the affected subjects will be generated:**
    - **Adverse Events**
    - **Concomitant Medications**
    - **Laboratory Data**

11. Literature Citations / References

Anderson CA, Boucher G, Lees CW, et al. Meta-analysis identifies 29 additive ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet*. 2011;43:246-252.

Danese S, Fiocchi C. Ulcerative Colitis. *N Engl J Med.* 2011;365(18):1713-1725.

Danese S, Fiorino G, Peyrin-Biroulet L, et al. Biological agents for moderately to severely active ulcerative colitis: a systematic review and network meta-analysis. *Ann Intern Med.* 2014;160(10):704-711.

Efavaleukin alfa Investigator's Brochure. Thousand Oaks, CA. Amgen Inc.

Geboes K, Riddell R, Öst A, Jensfelt B, Persson, T, Löfberg R. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. *Gut.* 2000;47 :404-409.

Malek TR, Castro I. Interlekin-2 Receptor Signaling: At the Interface between Tolerance and Immunity. *Immunity*. 2010;33(2):153-165.

Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative Colitis in Adults. *Am J Gastroenterol.* 2019;114(3):384-413.

Saadoun D, Rosenzwajg M, Joly F, et al. Regulatory T cell responses to low-dose interleukin-2 in HCV-induced vasculitis. *N Engl J Med.* 2011;365;2067-2077.

Sandborn WJ, Su C, Sands BE, et al. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. *N Engl J Med.* 2017;376(18): 1723-1736.

12. Appendices

Appendix A. Reference Values/Toxicity Grades

Assessment of severity for each adverse event and serious adverse event reported during the study will be based on:

The Common Terminology Criteria for Adverse Events (CTCAE), version 5 which is available at the following location:

http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm

Appendix B. Analytical Windows

Since the actual visit for a subject may not exactly coincide with their scheduled visit date, to allow for variations in scheduling, the following visit windows will be used to assign evaluations to a most appropriate nominal visit for analysis and summarization. Furthermore, there will be no gaps between visit windows in order to include as many data points as possible for summarization.

If more than one actual visit (including the unscheduled visits) falls within the same defined window, the visit closest to the target day with non-missing data will be considered for analysis. If two actual visit dates are at the same distance from the target day, the latest visit with non-missing data will be considered for analysis.

If more than one evaluation falls on the same date and time for laboratory results, then the value with the smallest accession number will be used.

Analysis visit windows for selected assessments are included in tables below. For remaining assessments, no visit window will be applied.

		Study Day Window		
Analysis Visit	**Target Day**	**Vital Signs**	**Chemistry, Hematology Urinalysis**	**Fecal Calprotectin**
Day 1	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]	2 - 70	
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]		
[redacted]	[redacted]	[redacted]	71 - 329	2 - 329

[redacted]	[redacted]		
		330 - 490	
Week 52			330 - 699
[redacted]	[redacted]		
		491 - 666	
		667 - 725	
Week 104 /ET		726 - 736	700 - 736

		Study Day Window

Analysis Visit	Target Day	Anti-efavaleukin alfa antibody	Lymphocyte subset	Partial Mayo Score, Physician Global Assessment (PGA)	Endoscopy, Geboes Score
Day 1	[redacted]				
[redacted]					
				2 - 35	
				36 - 63	
				64 - 105	
				106 - 119	
				120 - 147	
		2 - 329	2 - 329	148 -175	
				176 - 203	
				204 - 231	
				232 - 259	
				260 - 288	
				289 - 315	
				316 - 343	
Week 52		330 - 500	330 - 699	344 - 371	2= - 370
[redacted]				372 - 398	
				399 - 430	

[redacted]	[redacted]			431 - 458	
				459 - 490	
				491 - 521	
		501 - 725		522 - 552	
				553 - 583	
				584 - 610	
				611 - 635	
				636 - 666	
				667 - 697	
				698 - 709	
Week 104 /ET*		726 - 736	700 - 736	710 - 736	371 - 736

Appendix C. Handling of Dates, Incomplete Dates and Missing Dates

Imputation Rules for Partial or Missing Start Dates

The imputation Rules for Partial or Missing Start Dates for Medical History, Procedures, AE, Conmeds (except corticosteroid recorded ended due to steroid tapering), Hospitalization dates.

Start Date		Stop Date						
		Complete: *yyyymmdd*		Partial: *yyyymm*		Partial: *yyyy*		Missing
		< 1st dose	≥ 1st dose	< 1st dose *yyyymm*	≥ 1st dose *yyyymm*	< 1st dose *yyyy*	≥ 1st dose *yyyy*	
Partial: *yyyymm*	= 1st dose *yyyymm*	2	1	n/a	1	n/a	1	1
	≠ 1st dose *yyyymm*		2	2	2	2	2	2
Partial: *yyyy*	= 1st dose *yyyy*	3	1	3	1	n/a	1	1
	≠ 1st dose *yyyy*		3		3	3	3	3
Missing		4	1	4	1	4	1	1

1=Impute the date of first dose; 2=Impute the first of the month; 3=Impute January 1 of the year; 4=Impute January 1 of the stop year
Note: For subjects who were never treated (first dose date is missing), partial start dates will be set to the first day of the partial month or first day of year if month is also missing.

If stop date is partial, please follow below imputation

1. If date is missing, impute to end of the month
2. If month is missing, impute to end of the year, Dec 31st
3. If year is missing, event/medication is ongoing

Appendix D. Mayo Score Categories

Stool Frequency Subscore	**Score**
Normal number of stools for subject	0
1 or 2 stools more than normal	1
3 to 4 stools more than normal	2
5 or more stools than normal	3
Rectal Bleeding Subscore	**Score**
No blood seen	0
Streaks of blood with stool less than half of the time	1
Obvious blood (more than just streaks) or streaks of blood with stool most of the time	2
Blood alone passed	3
Endoscopic Subscore	**Score**
Normal or inactive disease	0
Mild disease (erythema, decreased vascular pattern)	1
Moderate disease (marked erythema, absent vascular pattern, friability, erosions)	2
Severe disease (spontaneous bleeding, ulceration)	3
Physician's Global Assessment	**Score**
Normal	0
Mild disease	1
Moderate disease	2
Severe disease	3
Mayo Score = Stool Frequency + Rectal bleeding+ Endoscopic Subscore + Physician's Global Assessment	

Note: The Mayo score range from 0 to 12, with higher score of indicating more severe disease. Modified Mayo score excluded Physician's Global Assessment and ranges from 0 to 9. Composite stool frequency and rectal bleeding score ranges from 0 to 6. The original description of the Mayo score included friability in the definition of an endoscopic subscore of 1. Consistent with current clinical practice and regulatory guidance, the study excluded friability from the definition of an endoscopic subscore of 1.

Stool frequency reports the number of stools in a 24-hour period, relative to the normal number of stools for that patient. The reference "normal" stool frequency for that patient will be recorded electronically at the screening visit. The Normal SF (stool frequency) refers to when the patient was in remission or, if the patient has never achieved remission, the reported stool frequency before initial onset of signs and symptoms of ulcerative colitis. Remission refers to a period of time since being diagnosed with ulcerative colitis when the patient is not experiencing any signs or symptoms relating to ulcerative colitis. This period of time may last a few weeks or a few months or may even last several years. The patient will record this electronically as souse data in the tablet device at the screening study visit.

Appendix E. Visit Date Considerations

- If the visit includes an actual endoscopy performed, then use the bowel preparation date.

- If the visit includes an actual endoscopy performed, but not getting 3 records from the bowel preparation date then use PGA assessment date.

- If the visit does not include an actual endoscopy performed, then use the PGA assessment date.

- If the visit does not include an actual endoscopy performed, but not getting 3 records from the PGA assessment date then use target assessment date.

- If the visit does not include an actual endoscopy performed and PGA assessment are also missing, then use target assessment date.

- If all above scenarios fails, then use worst case imputation.

The target assessment date for that visit should be used per Appendix B .

For sites located in France, site number 25001, 25002, 25003, 25004, and 25005, if a subject has an Endoscopy date recorded with a missing Bowel preparation date, then Endoscopy date will be considered for the Bowel preparation date. Bowel prep date may be missing as listed French sites may perform enema on the same day as endoscopy date. No imputation will be performed if Bowel preparation date is present.