Statistical Analysis Plan for Study M14-327

A Phase 2, Multicenter, Open-Label Extension (OLE) Study to Observe the Long-Term Efficacy, Safety, and Tolerability of Repeated Administration of Upadacitinib (ABT-494) in Subjects with Crohn's Disease

Version 1.0

Table of Contents

1.0 Introduction 4
2.0 Study Objectives and Design 4
2.1 Study Objectives 4
2.2 Study Design Overview 5
2.3 Treatment Assignment and Blinding 8
2.4 Sample Size Determination 8
3.0 Endpoints 9
3.1 Efficacy Endpoints 10
3.2 Safety Endpoints 12
4.0 Analysis Populations 12
5.0 Subject Disposition 13
6.0 Study Treatment Duration and Compliance 13
7.0 Subject Characteristics 14
7.1 Demographics and Baseline Characteristics 14
7.2 Medical History and Prior and Concomitant Medications 16
7.3 Protocol Deviations 16
8.0 Handling of Potential Intercurrent Events for the Efficacy Endpoints 17
9.0 Efficacy Analyses 17
9.1 General Considerations 17
9.2 Handling of Missing Data 18
9.3 Efficacy Endpoint(s) and Analyses 18
9.4 Efficacy Subgroup Analyses 18
10.0 Safety Analyses 18
10.1 General Considerations 18
10.2 Adverse Events 19
10.2.1 Treatment-Emergent Adverse Events 19
10.2.2 Adverse Event Overview 19
10.2.3 Treatment-Emergent Adverse Events by SOC and/or PT 20
10.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure 21

10.2.5 Deaths, Serious Adverse Events, and Adverse Events Leading to Study Treatment Discontinuation 22
10.2.6 Adverse Events of Special Interest 22
10.3 Analysis of Laboratory Data 22
10.4 Analysis of Vital Signs 25
11.0 Other Analyses 25
11.1 Patient-Reported Outcomes 25
11.1.1 EuroQol 5 Dimensions Questionnaire (EQ-5D) 25
11.1.2 Work Productivity and Activity Impairment Questionnaire: Crohn's Disease (WPAI:CD) 26
12.0 Interim Analyses 26
13.0 Overall Type-I Error Control 26
14.0 Version History 27
15.0 References 27

List of Tables

Table 1. SAP Version History Summary 27

List of Figures

Figure 1. Study Schematic 6

List of Appendices

Appendix A. List of SAP Signatories 28
Appendix B. Definition of Adverse Events of Special Interest 29
Appendix C. Potentially Clinically Important Criteria for Safety Endpoints 31

1.0 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analyses for Upadacitinib Study M14-327, A Phase 2, Multicenter, Open-Label Extension (OLE) Study to Observe the Long-Term Efficacy, Safety, and Tolerability of Repeated Administration of Upadacitinib (ABT-494) in Subjects with Crohn's Disease.

The analyses of pharmacokinetic endpoints will not be covered in this SAP.

The SAP will not be updated in case of administrative changes or amendments to the protocol unless the changes impact the analysis.

Unless noted otherwise, all analyses will be performed using SAS Version 9.4 (SAS Institute Inc., Cary, NC 27513) or later.

2.0 Study Objectives and Design

2.1 Study Objectives

The objective of study is to observe the long-term efficacy, safety, and tolerability of repeated administration of upadacitinib in subjects with CD who completed Study M13-740.

There is no hypothesis testing in M14-327 study. Descriptive statistics will be provided with respect to each efficacy point.

The estimand corresponding to each of the binary endpoints (Section 3.1) is defined as follow:

- The proportion of subjects achieving the corresponding response/remission status at the specified time point regardless of initiation of CD-related medications, regardless of premature discontinuation of study drug, within each treatment group among subjects with Crohn's disease who completed Study M13-740.

The estimand corresponding to each of the continuous endpoints (Section 3.1) is defined as follow:

- The mean change from Baseline/Week 0 of the corresponding variables at the specified time point regardless of initiation of CD-related medications, regardless of premature discontinuation of study drug, within each treatment group among subjects with Crohn's disease who completed Study M13-740.

The estimand corresponding to occurrence of hospitalizations/surgeries endpoints (Section 3.1) is defined as follow:

- The exposure-adjusted incidence rate of the corresponding events within the specified time period regardless of initiation of CD-related medications, regardless of premature discontinuation of study drug, within each treatment group among subjects with Crohn's disease who completed Study M13-740.

The estimand corresponding to the time to first dose escalation endpoint (Section 3.1) is defined as follow:

- The median time to first dose escalation regardless of initiation of CD-related medications, regardless of premature discontinuation of study drug, among subjects with Crohn's disease who completed Study M13-740 and started Upadacitinib 15 mg QD when enrolled in Study M14-327.

2.2 Study Design Overview

Study M14-327 is a Phase 2, multicenter, OLE study which comprises a 96 month follow up period designed to observe the long-term efficacy, safety, and tolerability of upadacitinib. Approximately 210 adult subjects who completed Study M13-740 and who met all of the inclusion criteria and none of the exclusion criteria will be eligible to enroll into this study.

Subjects will be evaluated for entry into Study M14-327 at the final study visit (Week 52) or up to 10 days following the Last Visit (Week 52) of Study M13-740. All subjects will

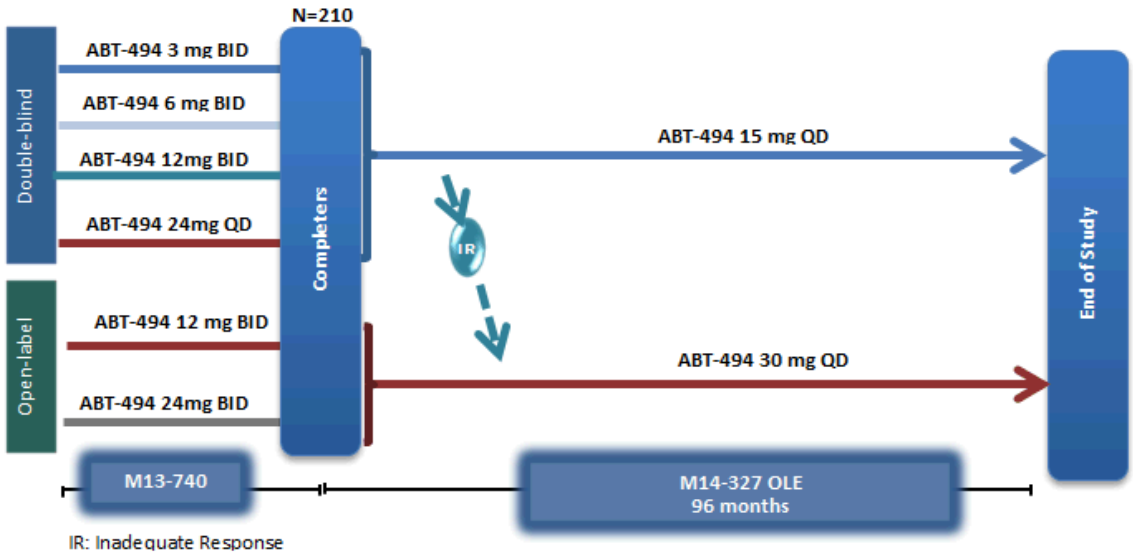
receive open-label upadacitinib beginning at Week 0. No dose will be administered at the Month 96 Visit.

A subject's participation in the study is anticipated to be up to 96 months. Study visits for clinical and safety assessments will be performed at Week 0, Month 1, Month 3, and then every 3 Months thereafter and until Month 96/Premature Discontinuation (PD). For study visit scheduling, a 30-day calendar (not 4 weeks) interval should be used to count each month. Endoscopic evaluation will be performed at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96/PD.

All subjects who were receiving the double-blind doses of upadacitinib during the extension phase of Study M13-740 will receive open-label upadacitinib 15 mg QD beginning at Week 0. The subjects who were receiving open-label doses (12 mg BID or 24 mg BID) during the extension phase of Study M13-740 will receive upadacitinib 30 mg QD in the Study M14-327.

The schematic of the study is shown in Figure 1.

Figure 1. Study Schematic



At or after Week 4, subjects receiving upadacitinib 15 mg QD who experience inadequate response and no safety concerns identified by the Investigator may be escalated to upadacitinib 30 mg QD.

At or after Week 4, subjects who initiated Study M14-327 with upadacitinib 30 mg QD (from the open-label doses in Study M13-740) who experience inadequate response may be withdrawn from the study, at the Investigator's discretion.

After at least 4 weeks of dose-escalation and treatment with 30 mg QD, subjects with persistent inadequate response may initiate or increase the dose of any CD-related medication to treat new or worsening of CD symptoms at the Investigator's discretion. The allowed CD-related medications are locally acting, oral, or intravenous corticosteroids, aminosalicylates, MTX, or CD-related antibiotics. AZA, 6-MP, cyclosporine, and tacrolimus are prohibited medications during the study and cannot be used as a rescue treatment. Changes in CD-related concomitant medications should be documented in the appropriate electronic case report form (eCRF).

After an additional 4 weeks, subjects who still meet the criteria for inadequate response may adjust (initiate or increase) the dose of any CD-related medication described above, at the Investigator's discretion. If after the adjustment of rescue therapy with the allowed CD-related medications, the subject continues to meet the criteria for inadequate response at the next study visit, the subject may be discontinued from the study.

Subjects taking 15 mg QD who dose escalate to 30 mg QD may have one opportunity to de-escalate to 15 mg QD provided the following criteria have been met:

- Be on current 30 mg QD for at least 3 months,
- Average daily liquid/soft Stool Frequency ≤ 1.5 and average daily Abdominal Pain score not worse than Baseline (of Study M14-327) OR average daily Abdominal Pain score ≤ 1.0 and average daily liquid/soft Stool Frequency not worse than Baseline (of Study M14-327), and
- Normal hs-CRP value (at the current or last visit) and fecal calprotectin < 150 μg/g at the current or last visit.

Subjects who experience inadequate response (using the same criteria outlined above) after dose de-escalation from upadacitinib 30 mg QD to 15 mg QD should be discontinued from the study.

Dose Recommendations for Subjects ≥ 65 Years

For Subjects 65 years and older receiving upadacitinib 30 mg QD, the Investigator will have the option to decrease the dose to upadacitinib 15 mg QD per Investigator discretion. For subjects who experience loss of efficacy and meet the criteria of inadequate response (as defined above) following a dose decrease from upadacitinib 30 mg QD to Upadacitinib 15 mg QD, the Investigator will have the option to escalate the dose to upadacitinib 30 mg QD once.

Dose de-escalation criteria for Subjects 65 years and older receiving upadacitinib 30 mg QD will be available once a site has obtained all required regulatory approvals and the related dose de-escalation functionality within the interactive response technology (IRT) system has been enabled.

2.3 Treatment Assignment and Blinding

All subjects who were receiving the double-blind doses of upadacitinib during the extension phase of Study M13-740 will receive open-label upadacitinib 15 mg QD beginning at Week 0. The subjects who were receiving open-label doses (12 mg BID or 24 mg BID) during the extension phase of Study M13-740 will receive upadacitinib 30 mg QD in the Study M14-327.

2.4 Sample Size Determination

No power calculation is conducted for Study M14-327 sample size. It is anticipated that up to 210 subjects will be enrolled in this OLE study.

3.0 Endpoints

The following endpoint definitions apply to the efficacy variables described below:

- **Remission:** Clinical Remission AND Endoscopic Remission
- **Response:** Clinical Response AND Endoscopic Response
- **Clinical Remission:** Average daily stool frequency ≤ 1.5 and not worse than Baseline AND average daily abdominal pain ≤ 1.0 and not worse than Baseline of Study M13-740
- **Modified Clinical Remission:** Average daily stool frequency ≤ 2.8 and not worse than Baseline AND average daily abdominal pain ≤ 1.0 and not worse than Baseline of Study M13-740, among subjects with stool frequency ≥ 4.0 or abdominal pain score ≥ 2.0 at Baseline of Study M13-740
- **Enhanced Clinical Response:** Average daily stool frequency at least 65% reduction from Baseline and average daily abdominal pain not worse than Baseline OR average daily abdominal pain at least 35% reduction from Baseline of Study M13-740 and average daily stool frequency not worse than Baseline of Study M13-740
- **Clinical Response:** Average daily stool frequency at least 30% reduction from Baseline and average daily abdominal pain not worse than Baseline OR average daily abdominal pain at least 30% reduction from Baseline of Study M13-740 and average daily stool frequency not worse than Baseline of Study M13-740
- **Endoscopic Remission:** SES-CD ≤ 4 and at least two-point reduction versus Baseline of Study M13-740 and no subscore > 1 in any individual variable
- **Endoscopic Improvement:** SES-CD at least 50% reduction from Baseline or endoscopic remission
- **Endoscopic Response:** Decrease ≥ 25% SES-CD from Baseline of Study M13-740
- **CDAI Remission:** CDAI < 150
- **CDAI Response:** CR-70: A reduction in CDAI by ≥ 70 from Baseline of Study M13-740

- **Enhanced CDAI Response:** CR-100: A reduction in CDAI by ≥ 100 from Baseline of Study M13-740
- **IBDQ Remission:** IBDQ ≥ 170
- **IBDQ Response:** Increase in IBDQ score ≥ 16 point from Baseline of Study M13-740

3.1 Efficacy Endpoints

- Achievement of:
  - remission at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
  - remission at Months 12, 24, 36, 48, 60, 72, 84, and 96 among subjects who achieved remission at Week 0
  - response at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
  - clinical remission at each visit where the variable is assessed according to scheduled study activities
  - modified clinical remission at each visit where the variable is assessed according to scheduled study activities
  - enhanced clinical response at each visit where the variable is assessed according to scheduled study activities
  - clinical response at each visit where the variable is assessed according to scheduled study activities
  - endoscopic remission at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
  - endoscopic remission at Months 12, 24, 36, 48, 60, 72, 84, and 96 among subjects who achieved endoscopic remission at Week 0
  - endoscopic improvement at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
  - endoscopic response at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
  - CDAI remission at each visit where the variable is assessed according to scheduled study activities
  - CDAI response at each visit where the variable is assessed according to scheduled study activities

- enhanced CDAI response at each visit where the variable is assessed according to scheduled study activities
- IBDQ remission at each visit where the variable is assessed according to scheduled study activities
- IBDQ response at each visit where the variable is assessed according to scheduled study activities
- steroid-free at each visit where the variable is assessed according to scheduled study activities, among subjects who took steroids at Baseline (of Study M13-740)
- steroid-free for at least 90 days and remission at Months 12, 24, 36, 48, 60, 72, 84, and 96, among subjects who took steroids at Baseline (of Study M13-740)
- remission and normal C-reactive protein (hs-CRP < 5 mg/L) at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
- remission and fecal calprotectin < 250 μg/g at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
- remission, and hs-CRP < 5 mg/L, and fecal calprotectin < 250 μg/g at Week 0, Months 12, 24, 36, 48, 60, 72, 84, and 96
- remission and discontinuation of corticosteroid uses at Months 12, 24, 36, 48, 60, 72, 84, and 96, among subjects who took corticosteroids at Week 0 of Study M13-740
- without draining fistulas at each visit where the variable is assessed according to scheduled study activities

- Time to first dose escalation among subjects who started upadacitinib 15 mg QD when enrolled in Study M14-327
- Occurrence of
  - CD-related hospitalization during study follow-up period
  - CD-related surgeries and procedures during study follow-up period
  - Dose escalation during study follow-up period, among subjects who started upadacitinib 15 mg QD when enrolled in Study M14-327

- Change from Baseline of Study M13-740 at each visit where the variable is assessed according to scheduled study activities in
  - average daily very soft/liquid stool frequency
  - average daily abdominal pain
  - CDAI
  - hs-CRP
  - fecal calprotectin
  - SES-CD
  - IBDQ
  - European Quality of Life 5 dimensions (EQ-5D-5L)
  - WPAI
- Changes in extraintestinal manifestations at each visit where the variable is assessed according to scheduled study activities

3.2 Safety Endpoints

The safety measures for this study include adverse event (AE) monitoring, physical examinations, vital sign measurements, and clinical laboratory testing (hematology, chemistry, and urinalysis).

4.0 Analysis Populations

The following population sets will be used for the analyses.

Intent-to-Treat (ITT) Population

ITT population: all subjects who enrolled into the study and received at least 1 dose of study drug. The ITT Population will be used for efficacy analyses.

Safety Analysis (SA) Population

SA population: all subjects who received at least 1 dose of study drug during the study. The SA Population will be used for safety analyses.

All analyses will be summarized by groups as follows:

- Group A: Subjects started Study M14-327 on 30 mg QD.
- Group B: Subjects started Study M14-327 on 15 mg QD and had dose escalated to 30 mg QD.
- Group C: Subjects started Study M14-327 on 15 mg QD and do not have dose escalated to 30 mg QD.

Of note, dose-escalation related efficacy endpoints will be summarized by combined Groups B and C (i.e., among all subjects who started upadacitinib 15 mg QD when enrolled in Study M14-327).

5.0 Subject Disposition

For the ITT population, a summary of subject accountability by investigator will be provided where the number of subjects in each of the following categories will be tabulated for each treatment group and overall:

- Subjects treated;
- Subjects who completed study treatment;
- Subjects who prematurely discontinued study treatment.

The number and percentage of subjects in the ITT population who prematurely discontinued study treatment will be summarized by reason for not completing study treatment overall and by treatment group.

6.0 Study Treatment Duration and Compliance

For the Safety Analysis Set, duration of treatment will be summarized for each treatment group and for total ABT-494 group. Duration of treatment is defined for each subject as last dose date minus first dose date + 1 day. Duration of treatment will be summarized using the number of subjects treated, mean, standard deviation, median, minimum and maximum. In addition, the number and percentage of subjects in each treatment duration

interval (≥ 1, ≥ 3, ≥ 6, ≥ 12, ≥ 18, ≥ 24, ≥ 36, ≥ 48, ≥ 60, ≥ 72, ≥ 84, ≥ 96 Months) will be summarized.

Treatment compliance will be summarized for the entire treatment period by treatment group and overall for the ITT population. Treatment compliance is defined as the number of tablets actually taken divided by the number of tablets that should have been taken. Percent compliance will be summarized.

7.0 Subject Characteristics

In Section 7.0, Baseline refers to the Baseline of Study M13-740. Categorical variables will be summarized with the number and percentage of subjects. Continuous variables will be summarized with descriptive statistics (number of non missing observations, mean and standard deviation, median, and minimum and maximum).

7.1 Demographics and Baseline Characteristics

Demographics and baseline or disease characteristics will be summarized descriptively, overall and by treatment group for the ITT population. Continuous demographic variables include age, weight, height, and body mass index (BMI), systolic blood pressure, diastolic blood pressure, pulse, and temperature.

Categorical demographic variables include:

- Sex
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)
- Race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Multi Race)
- Age (< 40, 40 – < 65, 65 – < 75, ≥ 75 years)
- Region (US or ex-US)
- Tobacco use (user, ex-user, never used, unknown)
- Alcohol use (drinker, ex-drinker, non-drinker, unknown)

The following disease characteristics and patient reported outcomes at both Baseline and Week 0 of Study M14-327 will be summarized:

- hs-CRP (mg/L)
- hs-CRP (< 10, ≥ 10 mg/L)
- Fecal Calprotectin Level (mcg/g)
- Fecal Calprotectin (≤ 250, > 250 mcg/g)
- CDAI
- CDAI (≤ 300, > 300)
- SES-CD
- SES-CD (< 15, ≥ 15)
- Duration of Crohn's Disease (years)
- Average Daily Stool Frequency
- Average Daily Abdominal Pain
- IBDQ – Total
- EuroQol-5D Index Score
- EuroQol-5D VAS
- WPAI – Work Time Missed (%)
- WPAI – Impairment While Working (%)
- WPAI – Overall Work Impairment (%)
- WPAI – Activity Impairment (%)

The following disease characteristics at Baseline will be summarized:

- Disease Duration (≤ 3, > 3 years)
- Baseline Corticosteroid use (yes, no)
- Baseline Immunosuppressant use (yes, no)
- Number of CD related prior anti-TNF agent use
- Number of CD related prior biologics use

7.2 Medical History and Prior and Concomitant Medications

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The actual version of the MedDRA coding dictionary will be noted in the statistical tables and clinical study report. The number and percentage of subjects in each medical history category (by MedDRA system organ class (SOC) and preferred term) will be summarized overall and by treatment groups. The SOC will be presented in alphabetical order, and the preferred terms will be presented in alphabetical order within each SOC. Subjects reporting more than one condition/diagnosis will be counted only once in each row (SOC or preferred term).

Prior and concomitant medications will be summarized separately. The number and percentage of subjects taking prior and concomitant medications will be summarized by generic drug name, based on the World Health Organization (WHO) Drug Dictionary. The actual version of the WHO Drug Dictionary will be noted in the statistical tables and clinical study report.

7.3 Protocol Deviations

For each of the following protocol deviation categories and across all categories, the number and percentage of randomized subjects with at least one protocol deviation will be summarized overall and by treatment group:

- Subject entered into the study even though did not satisfy entry criteria;
- Subject developed withdrawal criteria during the study but was not withdrawn;
- Subject received wrong treatment or incorrect dose of study treatment;
- Subject took excluded concomitant medication.

8.0 Handling of Potential Intercurrent Events for the Efficacy Endpoints

Potential intercurrent events considered in the Study M14-327 include (i) premature discontinuation of study drug (noted as ICE1 hereafter), (ii) initiation or dose increase of CD-related concomitant medications (noted as ICE2 hereafter).

Treatment policy strategy will be used to handle both ICE1 and ICE2. All data collected regardless of ICE1 and/or ICE2 will be included for all efficacy endpoints.

9.0 Efficacy Analyses

9.1 General Considerations

Descriptive statistics will be provided for all efficacy endpoints, by treatment groups among the ITT population. Dose-escalation related endpoints will be summarized among the combined Group B and C from the ITT population.

Categorical endpoints will be summarized by number and percentage of subjects who have achieved the endpoint, as well as the two-sided 95% confidence interval (CI; based on normal approximation) of the percentage. Continuous endpoints will be summarized by the number of subjects with observed data, mean (and its 95% CI), standard deviation, minimum, median, and maximum. Time to event endpoints will be summarized by sample size, number of events, and the 25%, 50%, and 75% percentiles of time (in days). Endpoints with respect to the occurrence of event will be summarized by sample size, number of events, exposure adjusted incidence rate (and its 95% CI based on Poisson distribution).

For efficacy analyses, Baseline is defined as the last non-missing observation measured on/before the day subjects received first dose of study drug in Study M13-740, and Week 0 of Study M14-327 is defined as the last non-missing observation measured on/before the day subjects received first dose of study drug in Study M14-327.

9.2 Handling of Missing Data

As observed (AO) analysis will be used for all efficacy analysis, i.e., missing evaluations will not be imputed. A subject who does not have an evaluation on a scheduled visit will be excluded from the AO analysis for that visit. AO will include all values collected in the study.

Last Observation Carried Forward (LOCF) approach will be used for efficacy endpoints that are measured frequently (e.g., within 6 months for two consecutive visits). The last available value from previous visits will be used to impute missing value for the same subject.

9.3 Efficacy Endpoint(s) and Analyses

All efficacy endpoints (as defined in Section 3.1) will be summarized by treatment groups among the ITT Population.

The attributes of the estimands corresponding to each of the efficacy endpoints are defined in Section 2.1.

9.4 Efficacy Subgroup Analyses

No efficacy subgroup analyses are planned in this study.

10.0 Safety Analyses

The Baseline for the analysis of laboratory data (Section 10.3) and vital sign data (Section 10.4) is defined as the last non-missing measurement before the first study drug administration in Study M13-740.

10.1 General Considerations

Safety data will be summarized among the SA population. Safety summaries will be presented by treatment group, including a total group for all subjects on active study treatment.

10.2 Adverse Events

Adverse events (AEs) will be summarized and presented using primary MedDRA System Organ Classes (SOCs) and preferred terms (PTs) according to the version of the MedDRA coding dictionary used for the study at the time of database lock. The actual version of the MedDRA coding dictionary used will be noted in the AE tables and in the clinical study report. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported.

10.2.1 Treatment-Emergent Adverse Events

Treatment-emergent adverse events (TEAEs) are defined as any AE with an onset after the first dose of study treatment in Study M14-327 and no more than 30 days after the last dose of study treatment. Events where the onset date is the same as the study treatment start date are assumed to be treatment-emergent. If an incomplete onset date is collected for an AE, then the AE will be assumed to be treatment-emergent unless there is evidence that confirms that the AE was not treatment-emergent (e.g., the known portion of the AE onset and/or the AE end date was prior to the date of the first dose of study treatment).

All treatment-emergent AEs will be summarized overall, as well as by primary MedDRA SOC and PT. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

The number and percentage of subjects experiencing treatment-emergent AEs will be summarized.

10.2.2 Adverse Event Overview

An overview of AEs will be presented consisting of the number and percentage of subjects experiencing at least one event for each of the following AE categories:

- Any TEAE

- Any TEAE related to study treatment according to the investigator
- Any TEAE leading to discontinuation of study treatment
- Any TEAE leading to death
- Any TEAE related to COVID-19
- Any severe TEAE
- Any serious TEAE
- Any Treatment-emergent AESI (Appendix B).
- All deaths
  - Deaths occurring ≤ 30 days after last dose of study drug
  - Deaths occurring > 30 days after last dose of study drug
  - Deaths related to COVID-19

In addition, an overview of AEs per 100 patient-years of study exposure will be presented for the AE categories defined above. The number of TEAEs reported, the total number of years of study drug exposure, and the TEAE rate per 100 patient-years will be presented.

10.2.3 Treatment-Emergent Adverse Events by SOC and/or PT

Treatment-emergent adverse events will be summarized by SOC and PT; by maximum relationship to study treatment as assessed by the investigator (e.g., reasonable possibility or no reasonable possibility) and SOC and PT; by maximum severity and SOC and PT; and by subject number and SOC and PT. Specific adverse events will be counted once for each subject for calculating percentages, unless stated otherwise. In addition, if the same adverse event occurs multiple times within a subject, the highest severity and level of relationship to investigational product will be reported. The SOCs will be presented in alphabetical order, and the PTs will be presented in alphabetical order within each SOC.

For summary by maximum severity, if a subject has an AE with an unknown severity, then the subject will be counted in the severity category of unknown, even if the subject has another occurrence of the same event with a severity present. The only exception is if

the subject has another occurrence of the same AE with the most extreme severity – severe. In this case, the subject will be counted under the severe category.

For summary by maximum relationship to study drug, if a subject has an AE with an unknown relationship, then the subject will be counted in the relationship category of "unknown," even if the subject has another occurrence of the same event with a relationship present. The only exception is if the subject has another occurrence of the same AE with a relationship assessment of "Reasonable Possibility." In this case, the subject will be counted under the "Reasonable Possibility" category.

In addition, treatment-emergent adverse events will be summarized by PT and sorted by decreasing frequency for the total active group.

10.2.4 Treatment-Emergent Adverse Events per Patient-Years of Exposure

TEAEs will be summarized by event rate per 100 subject years, defined as

$$100 \times \frac{Number\ of\ TEAEs}{Total\ Patient\ Years}$$

where total patient years is defined as the sum of the study drug exposure of all subjects, normalized by 365.25, and rounded to 1 decimal place.

Study drug exposure among the Safety Population is defined for each subject as the last dose date minus the first dose date +1.

Note that one event per preferred term per day per subject will be counted in the calculation of the number of AEs (i.e., a preferred term will not be counted twice on the same day for the same subject).

The TEAE rates per 100 patient-years of exposure will be provided for each AE category in the AE overview summary (defined in Section 10.2.1) and for TEAE summary by SOC and PT.

10.2.5 Deaths, Serious Adverse Events, and Adverse Events Leading to Study Treatment Discontinuation

Treatment-emergent serious adverse events (TESAEs), TEAEs leading to premature discontinuation of study treatment, and TEAEs leading to death will be summarized by SOC and PT.

Tabular listings will be provided for all deaths, all SAEs, TESAEs, TEAEs leading to death, TEAEs leading to premature discontinuation of study treatment, and TEAEs leading to study treatment interruptions.

10.2.6 Adverse Events of Special Interest

The AESI categories will be identified by the search criteria per Standard MedDRA Queries (SMQs)/Company MedDRA Queries (CMQs) specified in Appendix B.

Treatment-emergent AESIs will be summarized by SOC and PT and listing format. Additionally, AESI rates per 100 patient years of exposure will be provided for each AESI category in the AE overview summary.

10.3 Analysis of Laboratory Data

Data collected from central and local laboratories, including additional laboratory testing due to an SAE, will be used in all analyses, except for Baseline where SAE-related laboratory assessments on or before the first dose of study drug will be excluded. The clinical laboratory tests defined in the protocol (e.g., hematology, clinical chemistry, and urinalysis) will be summarized.

Analysis of Quantitative Laboratory Parameters

Each laboratory variable will be summarized for all time points (starting with Week 0 of Study M14-327) with visit mean. The change from Baseline mean, standard deviation, median, minimum and maximum will be presented for the mean change from Baseline within each treatment group.

Shift Table Analyses

Changes in laboratory parameters will be tabulated using shift tables categorized as low, normal, or high based on the normal ranges of the laboratory used for each sample. A shift table from Baseline to minimum and maximum value (based on normal range), will be created. A similar shift table will be provided to summarize shifts from Baseline to the final post-Baseline value.

Potentially Clinically Significant Laboratory Values

The criteria for potentially clinically significant laboratory values will be determined by CTCAE criteria of Grade 3, Grade 4 and ≥ Grade 3, with a grade worsening compared to [redacted] criteria for potentially clinically significant laboratory values will be summarized by treatment group among the SA1 Population. Listings of subject-level laboratory data will be provided for subjects meeting the criteria.

Assessment of Liver Enzyme Elevations

The liver-specific laboratory tests include the serum glutamic pyruvic transaminase (SGPT/ALT), serum glutamic-oxaloacetic transaminase (SGOT/AST), alkaline phosphatase (ALP), and total bilirubin (TBL). The frequencies and percentages of subjects with post Baseline liver specific function test values that meet the following criteria of potential clinical interest will be summarized by treatment group:

- ALT > 3 × ULN
- ALT > 5 × ULN
- ALT > 10 × ULN
- ALT > 20 × ULN
- AST > 3 × ULN
- AST > 5 × ULN
- AST > 10 × ULN
- AST > 20 × ULN
- TBL > 2 × ULN
- ALP > 1.5 × ULN
- ALT and/or AST > 3 × ULN and TBL > 1.5 × ULN
- ALT and/or AST > 3 × ULN and TBL > 2 × ULN,

where ULN is the upper normal limit. The maximum ratio relative to the ULN will be used to determine if subjects meet any of the criteria listed above.

A listing of possible Hy's Law cases, defined as those meeting the following criteria will also be provided:

[redacted]

10.4 Analysis of Vital Signs

Vital sign measurements of systolic and diastolic blood pressure, pulse rate, respiratory rate, body weight, and body temperature will be summarized.

Each vital sign variable will be summarized for all time points (starting with Week 0 of Study M14-327) with visit mean. The change from Baseline mean, standard deviation, median, minimum and maximum will be presented for the mean change from Baseline within each treatment group.

Vital sign variables will be evaluated based on potentially clinically important (PCI) criteria (Appendix C). For each vital sign PCI criterion, the number and percentage of subjects who have a vital sign value meeting the criteria will be summarized. Listings will be provided to summarize subject-level vital sign data for subjects meeting PCI criteria.

11.0 Other Analyses

11.1 Patient-Reported Outcomes

The following patient-reported outcomes (PROs) are being collected: EuroQol-5 Dimensions-5 Level (EQ-5D-5L) and Work Productivity and Activity Impairment: Crohn's Disease (WPAI:CD).

11.1.1 EuroQol 5 Dimensions Questionnaire (EQ-5D)

EuroQol 5 Dimensions 5 Level Questionnaire (EQ-5D-5L)

The EQ-5D-5L has two parts: the EuroQol 5 Dimensions (EQ-5D) descriptive system and the EQ Visual Analogue Scale (EQ VAS). The EQ-5D-5L has five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. These dimensions are measured on a five-level scale: no problems, slight problems, moderate problems, severe problems, and extreme problems. The responses for the 5 dimensions are used to compute a score which will be converted into a single health utility index score.[1,2] The VAS score will be analyzed separately and is not included in scoring the dimensions or the utility

calculation. For the EQ VAS score, higher scores indicate better health-related quality of life. An increase from Baseline in the EQ-5D-5L health index score represents an improvement.

For the EQ-5D-5L health index and EQ VAS scores, no imputation will be performed for missing items. A complete health state (non-missing responses to each of the 5 dimensions) is needed to derive the EQ-5D health index score.

Descriptive, summary statistics at each visit and for change from Baseline to each visit will be provided for the EQ-5D-5L health index and VAS scores by treatment group.

11.1.2 Work Productivity and Activity Impairment Questionnaire: Crohn's Disease (WPAI:CD)

The WPAI:CD includes 6 questions and will be used to determine work productivity and impairment levels due to Crohn's Disease. The WPAI can yield 4 types of scores: absenteeism (work time missed), presenteeism (impairment at work/reduced on-the-job effectiveness), work productivity loss (overall work impairment), and activity impairment. WPAI outcomes are expressed as impairment percentages, with higher percentages indicating greater impairment and less productivity, i.e., worse outcomes.

Descriptive, summary statistics at each visit and for change from Baseline to each visit will be provided for the absenteeism, presenteeism, activity impairment, and overall work impairment scores by treatment group.

12.0 Interim Analyses

No formal interim analysis is planned for this study.

13.0 Overall Type-I Error Control

There is no hypothesis testing in this open-label extension study. Accordingly, there is no overall type-I error control.

14.0 Version History

Table 1. SAP Version History Summary

Version	Date	Summary
1.0	28 February 2025	Initial version

15.0 References

1. EuroQol Research Foundation. EQ-5D-5L user guide. 2019. Available from: https://euroqol.org/.
2. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) measurement system: properties, applications, and interpretation. Health Qual Life Outcomes. 2003;1:79.

Appendix A. List of SAP Signatories

Name	Title	Role/Functional Area
[redacted]	[redacted]	Therapeutic Area Medical Director
		Project Lead Programmer
		Statistics TA Lead
		Director of Statistics
		Author

Appendix B. Definition of Adverse Events of Special Interest

AESI will be identified by the following CMQ, SMQ, and other search criteria:

Table B-1. AESI for Upadacitinib with SMQs/CMQs/PTs Searches

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Serious Infections	[redacted]		
Opportunistic Infections excluding Tuberculosis and Herpes Zoster			
Active Tuberculosis			
Herpes Zoster			
Adjudicated Gastrointestinal Perforations			
Anemia			
Neutropenia			
Lymphopenia			
Creatine Phosphokinase (CPK) Elevations			
Hepatic Disorders			
Renal Dysfunction			
Malignancies (all types)			
Malignancies excluding non-melanoma skin cancer (NMSC)			
NMSC			

AESI	Type of MedDRA Query	Broad or Narrow Search	SMQ/CMQ Search Criteria
Lymphoma	[redacted]		
Adjudicated Cardiovascular Events			
MACE*			
Cardiovascular Death			
Non-fatal Myocardial Infarction			
Non-fatal Stroke			
Other Cardiovascular Events			
Undetermined/Unknown Cause of Death			
Adjudicated Thrombotic Events			
VTE**			
Deep Vein Thrombosis			
Pulmonary Embolism			
Other Venous Thrombosis			
Arterial Thromboembolic Events (non-cardiac, non-neurologic)			

CAC = Cardiovascular Adjudication Committee; CMQ = company MedDRA query; PT = preferred term; SMQ = standard MedDRA query

* MACE: Major Adverse Cardiovascular Events, defined as cardiovascular death, non-fatal myocardial infarction and non-fatal stroke.

** VTE: Venous thromboembolic events, defined as deep vein thrombosis (DVT) and pulmonary embolism (PE) (fatal and non-fatal).

a. Reviewed and adjudicated by an independent Cardiovascular Adjudication Committee in a blinded manner.

Appendix C. Potentially Clinically Important Criteria for Safety Endpoints

The criteria for Potentially Clinically Important (PCI) criteria for vital sign findings are described in Table C-1.

Table C-1. Criteria for Potentially Clinically Important Vital Sign Values

Vital Sign	Category	Criteria for Potential Clinically Significant Vital Signs
Systolic blood pressure	Low	[redacted]
	High	
Diastolic blood pressure	Low	
	High	
Pulse	Low	
	High	
Weight	High	
	Low	

Document Approval

Study M14327 - Statistical Analysis Plan Version 1 - 28Feb2025 (E3 16.1.9)

Version: 1.0 **Date:** 03-Mar-2025 **Company ID:** 20250303-0900f9f689299d54-1.0-en

Signed by:	Date:	Meaning of Signature:
[redacted]	03-Mar-2025 16:37 UTC	Approver - Statistics
	03-Mar-2025 02:36 UTC	Approver
	28-Feb-2025 17:22 UTC	Approver
	28-Feb-2025 17:04 UTC	Approver - Statistics
	28-Feb-2025 16:56 UTC	Approver - Statistics