Statistical Analysis Plan

A 14-Week Phase 2b, RandomizEd, Double-BLind, Dose-Ranging Study to Determine the PharmacokInetics, Efficacy, Safety, and Tolerability of TEV-48574 in Adult PatiEnts with Moderate to Severe Ulcerative Colitis or Crohn's Disease (RELIEVE UCCD)

Study Number TV48574-IMM-20036

NCT05499130

SAP Approval Date: 22 October 2024

Statistical Analysis Plan with Amendment 02

Study TV48574-IMM-20036

A 14-Week Phase 2b, RandomizEd, Double-BLind, Dose-Ranging Study to Determine the PharmacokInetics, Efficacy, Safety, and Tolerability of TEV-48574 in Adult PatiEnts with Moderate to Severe Ulcerative Colitis or Crohn's Disease (RELIEVE UCCD)

Short title: **A Randomized, Double-Blind Study on Pharmacokinetics, Efficacy, Safety, and Tolerability of TEV-48574 in Adults with IBD**

Lay title: **A Study to Test the Effect of TEV-48574 in Moderate to Severe Ulcerative Colitis or Crohn's Disease**

Dose-Ranging Study (Phase 2b)

IND number: 157634; NDA number: Not applicable; BLA number: Not applicable; EU CT number: 2024-511089-36-00

Version: Final with Amendment 02

Approval Date: 22 October 2024

Sponsor

Teva Branded Pharmaceutical Products R&D, Inc.
145 Brandywine Parkway
West Chester, Pennsylvania 19380
United States of America

Prepared by: [redacted]

[redacted]

STATISTICAL ANALYSIS PLAN APPROVAL

Study No.: TV48574-IMM-20036

Study Title: A 14-Week Phase 2b, RandomizEd, Double-BLind, Dose-Ranging Study to Determine the PharmacokInetics, Efficacy, Safety, and Tolerability of TEV-48574 in Adult PatiEnts with Moderate to Severe Ulcerative Colitis or Crohn's Disease (RELIEVE UCCD)

Statistical Analysis Plan for:

☒ Interim Analysis ☐ Integrated Summary of Efficacy

☒ Final Analysis ☐ Integrated Summary of Safety

Version: Final with Amendment 02

Author: [redacted]

Approver: [redacted]

Approver: [redacted]

Executed signature pages are maintained separately within the Trial Master File

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AMENDMENT HISTORY

The Statistical Analysis Plan (SAP) for trial TV48574-IMM-20036 has been amended based on the changes in Protocol Amendment 04.

A comparison of the major changes the primary analysis model between Protocol Amendment 02 (original SAP) and Protocol Amendment 04 (SAP amendment 02), with justification, is presented in Appendix 3.

Amendment number	Date	Author(s)	Summary of major changes	Reason for amendment
01	11 Mar 2024	[redacted]	Changes according to Protocol Amendment 03 with Revision 01	Protocol Amendment
02	22 Oct 2024	[redacted]	Interim analysis for efficacy/futility will not be performed	The efficacy /futility interim analysis will be impractical and have limited utility, as all participants have been randomized at the time that the data cut for this interim analysis would have taken place.
			Supplementary estimand for the primary estimand updated to include for UC clinical remission based on MMS at week14 a requirement that the stool frequency subscore is not greater than baseline.	This addition was done to align with the definition of clinical remission (MMS) currently used in clinical development of therapies for UC.
			Clarification regarding intercurrent events	Clarification
			Exploratory efficacy endpoints added.	Addition to protocol-defined exploratory efficacy endpoints
			Pharmacokinetic (PK) analysis section added; definition of PK analyses set added.	Addition of descriptive statistics of serum concentrations of TEV-48574 to the SAP

Amendment number	Date	Author(s)	Summary of major changes	Reason for amendment
			Immunogenicity analysis section revised.	Clarification.
			Appendices for derivation of CDAI score and PRO-CD Score added.	Clarification

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
ADA	anti-drug antibody
AE	adverse event
BL	baseline
BMI	body mass index
CD	Crohn's Disease
CDAI	Crohn's Disease Activity Index
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
ECG	Electrocardiogram
EU	European Union
ET	early termination
FSH	follicle stimulating hormone
GI	gastrointestinal
HCT	hematocrit
hsCRP	high sensitivity C-reactive protein
IA	Interim Analysis
IBD	inflammatory bowel disease
IDMC	Independent Data Monitoring Committee
ITT	Intent-to-Treat
IMP	investigational medicinal product
JAK	Janus kinase
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intent-to-Treat
MMS	modified Mayo Score
(MM) SES-CD	modified multiplier Simple Endoscopic Score for Crohn's Disease
NRS	numerical response scale
[redacted]	[redacted]
PD	Pharmacodynamic
PK	Pharmacokinetic
PRO2	2-item patient-reported outcome

Abbreviation	Term
PT	preferred term
Q2W	every 2 weeks
R&D	Research and Development
RTSM	Randomization and Trial Supply Management
RNA	ribonucleic acid
S1P	sphingosine-1-phosphate
SAP	Statistical Analysis Plan
sc	subcutaneous
SD	Standard Deviation
SE	Standard Error
SES-CD	Simple Endoscopic Score for Crohn's Disease
SI	standard international
SOC	system organ class
SOP	standard operating procedure
TL1A	tumor necrosis factor (ligand) 1A
TNF-α	tumor necrosis factor-alpha
UC	Ulcerative Colitis
ULN	upper limit of normal
WHO	World Health Organization

INTRODUCTION

This Statistical Analysis Plan (SAP) describes the planned analysis and reporting for Teva Branded Pharmaceutical Products R&D, Inc. study TV48574-IMM-20036, (A 14-Week Phase 2b, RandomizEd, Double-BLind, Dose-Ranging Study to Determine the PharmacokInetics, Efficacy, Safety, and Tolerability of TEV-48574 in Adult PatiEnts with Moderate to Severe Ulcerative Colitis or Crohn's Disease (RELIEVE UCCD)), and was written in accordance with SOP GSD_RD_702 (Statistical Analysis Plan).

The reader of this SAP is encouraged to read the study protocol for details on the conduct of this trial, the operational aspects of clinical assessments, and the timing for completing the participation of a participant in this trial.

The SAP is intended to be in agreement with the protocol, especially with regards to the primary and all secondary endpoints and their respective analyses. However, the SAP may contain more details regarding these particular points of interest, or other types of analyses (e.g., other endpoints). When differences exist in descriptions or explanations provided in the study protocol and this Statistical Analysis Plan, the Statistical Analysis Plan prevails; the differences will be explained in the Clinical Study Report (CSR).

1. TRIAL OBJECTIVES AND ENDPOINTS

1.1. Primary and Secondary Trial Objectives and Endpoints

The primary and secondary trial objectives and endpoints are:

Objectives	Endpoints
The **primary objective** of the trial is to characterize the efficacy of TEV-48574 sc administered Q2W in adult participants with IBD (IBD with moderate to severe UC or CD), as assessed by induction of clinical remission (UC) and endoscopic response (CD) at week 14.	The primary efficacy endpoint is response (yes or no) at week 14, where response in UC participants is defined as clinical remission and response in CD participants is defined as endoscopic response. Clinical remission and endoscopic response are determined as follows: • Clinical remission at week 14 in participants with moderate to severe UC. Clinical remission is a modified (9-point rectal bleeding, stool frequency, and endoscopy) Mayo score (MMS) of ≤2 points, which is defined by: – stool frequency subscore of 0 or 1, – rectal bleeding subscore of 0, and – endoscopic subscore of 0 or 1, where a score of 1 does not include "friability" • Endoscopic response at week 14 in participants with moderate to severe CD, defined as a reduction in Simple Endoscopic Score for Crohn's Disease (SES--CD) of at least 50% from baseline
A **secondary objective** of the trial is to evaluate the efficacy and dose response of 2 different dose regimens of TEV-48574 sc administered Q2W in adult participants with IBD (moderate to severe UC or CD) as assessed by multiple standard measures at week 14.	The secondary efficacy endpoints to be measured in participants with moderate to severe UC are as follows: • Clinical response at week 14, defined as a decrease from baseline in MMS of at least 2 points AND at least a 30% reduction from baseline with either a decrease in rectal bleeding subscore of at least 1 or an absolute rectal bleeding subscore of less than or equal to 1 • Endoscopic improvement defined as a Mayo endoscopic subscore of 0 or 1, where a score of 1 does not include "friability", at week 14 • Endoscopic remission defined as a Mayo endoscopic subscore of 0 at week 14 • Clinical response defined as decrease from baseline of at least 50% in 2-item patient-reported outcome (PRO2-UC; rectal bleeding and stool frequency) at week 14

Objectives	Endpoints
	• Clinical remission defined as score of rectal bleeding = 0 and stool frequency = 0 on the PRO2-UC scale at week 14 The secondary efficacy endpoints to be measured in participants with moderate to severe CD are as follows: • Clinical response defined as a ≥100-point decrease in Crohn's Disease Activity Index (CDAI) score from baseline at weeks 4, 8, 12 and 14 • Clinical remission defined as a CDAI score <150 at week 14 • Clinical response defined as a decrease from baseline of at least 50% in PRO2 (PRO2-CD; abdominal pain and stool frequency) at week 14 • Clinical remission defined as abdominal pain ≤1 and stool frequency ≤3 on the PRO2 scale at week 14 • Endoscopic response defined as a decrease in modified multiplier (MM)-SES-CD of >50% from baseline at week 14
A **secondary objective** of the trial is to evaluate the safety and tolerability of 3 different dose regimens of TEV-48574 sc administered Q2W in adult participants with IBD (moderate to severe UC or CD).	The safety and tolerability measures/parameters are as follows: • Adverse events • Change from baseline in clinical laboratory test results (serum chemistry, hematology, and urinalysis) • Change from baseline in vital signs measurements (blood pressure, pulse rate, body temperature, and respiratory rate) • Change from baseline in 12-lead electrocardiogram findings • Participants who stopped the investigational medicinal product due to adverse events • Local tolerability at the injection site
A **secondary objective** of the trial is to evaluate the immunogenicity of 3 different dose regimens of TEV-48574 sc administered Q2W in adult participants with IBD (moderate to severe UC or CD).	The immunogenicity endpoints for this trial are as follows: • Treatment-emergent anti-drug antibody (ADA) at weeks 2, 4, 8, 14, and follow-up visit

Objectives	Endpoints
	• Presence of neutralizing ADA in ADA positive participants at weeks 2, 4, 8, 14, and follow-up visit

1.2. Primary Estimand

The primary estimand in this trial is defined by the following attributes:

- **Treatment:** In both UC and CD groups, participants will be randomly assigned to receive any one of the trial treatment regimens shown below, in a 1:1:1 ratio:
  - TEV-48574 2250 mg (single loading dose)/900 mg (6 induction doses)
  - TEV-48574 2250 mg (single loading dose)/450 mg (6 induction doses)
  - Matching placebo
- **Target population:** The target population comprises participants with a diagnosis of moderate to severe UC or moderate to severe CD.
- **Variable:** Response for the primary efficacy analysis is defined as clinical remission status at week 14 for UC participants or endoscopic response status at week 14 for CD participants.
- **Handling of intercurrent events:**

Intercurrent Event	Policy	Action
Participant did not receive at least 1 dose of investigational medicinal product (IMP)	Principal strategy	Participant will be excluded from the analysis set as per modified intent-to-treat (mITT) definition
Use of rescue medication, including increase in dose of systemic corticosteroids for IBD [a]	Composite strategy	Participant will be reported as non-remitter/non-responder
Major UC/CD related surgery [b]		
Treatment discontinuation (regardless of the reason)		
Early termination during the treatment period (regardless of the reason)		
Important protocol deviations, including use or changes in other medications [a]	Treatment policy	Data after the intercurrent event will be used

[a] See Section 4.5 for classification of initiation or dose increase of IBD medications
[b] See Section 4.6 for classification of IBD-related surgery

- **Population-level summary:** The population-level summary of interest is the posterior distribution $(p_{d,i} - p_{0,i})$ estimated from a Bayesian Beta-Binomial model for the response rate (clinical remission rate for UC, or endoscopic response rate for CD), $p_{d,i}$, where i=*UC, CD* and d =0,1,2 represent placebo, TEV-48574 450 mg Q2W, and TEV-48574 900 mg Q2W, respectively.

1.2.1. Supplementary Estimands

The first supplementary estimand is defined by the following attributes:

- **Treatment**: same as for the Primary Estimand.
- **Target population**: same as for the Primary Estimand.
- **Variable**:
  - UC: clinical remission status at week 14, where clinical remission is defined as MMS ≤2 points, with stool frequency subscore of 0 or 1 and not greater than baseline, rectal bleeding subscore of 0, and endoscopic subscore of 0 or 1, where a score of 1 does not include "friability".
  - CD: endoscopic response status at week 14, where endoscopic response is defined as a reduction in SES--CD of > 50% from baseline or a decrease of at least 2 points for participants with a baseline score of 4 and isolated ileal disease.
- **Handling of intercurrent events**: same as for the Primary Estimand.
- **Population-level summary**: same as for the Primary Estimand.

The second supplementary estimand is defined by the following attributes:

- **Treatment**: same as for the Primary Estimand.
- **Target population**: same as for the Primary Estimand.
- **Variable**: same as for the Primary Estimand.
- **Handling of intercurrent events**: same as for the Primary Estimand.
- **Population-level summary**: The population-level summary is the difference in response rates of each TEV-48574 dose compared to placebo obtained from a logistic regression model fit separately to each indication, with dose as categorical variable.

1.3. Exploratory Objectives and Endpoints

The exploratory objectives and endpoints are:

Objectives	Endpoints
An **exploratory objective** of the trial is to evaluate the efficacy of 2 different doses of TEV-48574 sc administered Q2W in adult participants with IBD (moderate to severe UC or CD) as assessed by multiple standard measures.	The exploratory efficacy endpoints to be measured in participants with moderate to severe UC are as follows: • Clinical response defined as decrease from baseline of at least 50% in 2-item patient-reported outcome (PRO2; rectal bleeding and stool frequency) at weeks 2, 4, 6, 8, 10, and 12 • Clinical remission defined as score of rectal bleeding = 0 and stool frequency = 0 on the PRO2 scale at weeks 2, 4, 6, 8, 10, and 12 • Histological remission defined as a Robarts Histopathology Index of ≤5 at week 14 • Histological remission defined as Geboes index score ≤3.1 at week 14 • Histologic-Endoscopic Mucosal Improvement (HEMI) defined as a Mayo endoscopic subscore of 0 or 1 without evidence of friability and Geboes score ≤3.1 The exploratory efficacy endpoints to be measured in participants with moderate to severe CD are as follows: • Clinical response defined as a decrease from baseline of at least 50% in PRO2 (PRO2 is defined as having 2 components, abdominal pain and stool frequency) at weeks 2, 4, 6, 8, 10, and 12 • Clinical remission defined as abdominal pain ≤1 and stool frequency ≤3 on the PRO2 scale at weeks 2, 4, 6, 8, 10, and 12 • Clinical remission defined as abdominal pain ≤1 and stool frequency ≤1.5 on the PRO2 scale at week 14 • Endoscopic remission defined as SES CD score of 0-2 • Endoscopic remission defined as SES CD score of 0-4,, with no individual subscore >1 at week 14 • Histologic response defined as a ≥50% decrease in Global Histologic Activity Score from baseline at week 14

Objectives	Endpoints
An **exploratory objective** of the trial is to evaluate the safety and tolerability of 2 different doses of TEV-48574 sc administered Q2W in adult participants with IBD (moderate to severe UC or CD).	• Use of concomitant medication • Device-related adverse events and malfunctions (for the commercial sc infusion system)
An **exploratory objective** of this trial is to evaluate association among exploratory biomarkers and clinical efficacy of TEV-48574 in adult participants with IBD (moderate to severe UC or CD).	• Change from baseline at weeks 2, 4, 8, and 14 in protocol specified- serum-resident pharmacodynamic (PD) markers • Change from baseline in serum and/or gastrointestinal (GI) tissue markers of GI tissue condition (at weeks 2, 4, 8, and 14 for serum; at week 14 for tissue) • Change from baseline at week 14 in GI tissue transcriptome • Change from baseline in fecal calprotectin at weeks 2, 4, 8, and 14 • Change from baseline in high sensitivity C-reactive protein (hsCRP) at weeks 2, 4, 8, and 14 • Change from baseline in blood albumin levels at weeks 2, 4, 8, and 14 • Change from baseline in serum free and total tumor necrosis factor (ligand) 1A at weeks 2, 4, 8, and 14 • Change from baseline at week 14 in GI tissue tumor necrosis factor (ligand) 1A (TL1A) expression • Change from baseline in UC-100 at week 14 in participants with moderate to severe UC • Change from baseline in [redacted] at week 2, 8, and 14

Objectives	Endpoints
An **exploratory objective** of the trial is to obtain trough serum TEV-48574 concentrations, to compare major pharmacokinetic (PK) characteristics between UC and CD participants with healthy volunteers and asthma participants, and, if data allows, to evaluate the pharmacokinetics/pharmacodynamics and/or exposure-response relationship of 3 different dose regimens of TEV-48574 sc.	• Trough serum TEV-48574 concentrations throughout the trial (sparse sampling) • Population pharmacokinetic parameters (e.g., clearance, volume of the central compartment, area under the concentration-time curve, maximum observed drug concentration, and trough drug concentration) • Population pharmacokinetic-pharmacodynamic parameters • Exposure-response parameters
An **exploratory objective** of this trial is to evaluate the effect of genetic polymorphisms on clinical efficacy in adult participants with IBD (moderate to severe UC or CD).	1. Primary and other efficacy endpoints

In addition, the following efficacy endpoints are considered as exploratory endpoints:

- PRO2-UC items at weeks 2, 4, 6, 8, 10, 12 and 14
- PRO2-UC total score at weeks 2, 4, 6, 8, 10, 12 and 14
- CDAI score at weeks 4, 8, 12 and 14
- CDAI remission at weeks 4, 8, and 12
- PRO2-CD remission at weeks 2, 4, 6, 8, 10, 12, and 14, defined as abdominal pain ≤1 and stool frequency ≤2.8 on the PRO2-CD scale.
- PRO2-CD items at weeks 2, 4, 6, 8, 10, 12 and 14
- PRO2-CD total score at weeks 2, 4, 6, 8, 10, 12 and 14
- Clinical response at weeks 2, 4, 6, 8, 10, and 12 defined as a decrease from baseline of at least 30% in PRO2-CD

2. TRIAL DESIGN

This section describes the trial design according to Protocol Amendment 03 Rev01. For details on the trial design prior to this protocol amendment, see Section 16 of the protocol.

2.1. General Design

This is a Phase 2b, randomized, double-blind, dose-ranging trial to determine the pharmacokinetics, efficacy, safety, and tolerability of TEV-48574 in adult participants with moderate to severe UC or CD. The trial will enroll adult participants (18 to 75 years of age, inclusive) of male and female sex (without restrictions on gender) with moderate to severe active UC or CD and who have demonstrated an inadequate response to, loss of response to, or intolerance to at least 1 of the following therapies: corticosteroids, immunosuppressant, or an advanced therapy for IBD including biologics, (anti-TNF, anti-integrins, anti-IL-12/23 or anti-IL-23), Janus kinase (JAK) inhibitors, or sphingosine-1-phosphate (S1P) receptor modulators, and no more than 3 locally approved classes of biologics (see protocol section 4.1 for definitions of these terms).

The trial will consist of a screening period of up to 6 weeks (42 days), a 14-week treatment period, and a 4-week follow-up period.

Screening: After providing written informed consent, participants will be screened within 6 weeks (42 days) prior to randomization to confirm that they have met all the selection criteria for the trial. At least 2 visits to the investigational site will be necessary to complete all screening procedures, including an endoscopy. The endoscopy should be performed after key eligibility criteria have been met and within approximately 10 calendar days of randomization (day 1) to allow for central endoscopy scoring.

Treatment and Follow-Up Period: During the 14-week treatment period, participants will visit the site Q2W on days 1, 15, 29, 43, 57, 71, and 85 (±3 days) for investigational medicinal product (IMP) administration (7 visits), as well as an end-of-treatment visit on day 99 (±3 days; week 14). After the end of the 14-week treatment period, all participants may be offered the option to enter a long-term extension trial described in a separate protocol (TV48574-IMM-20038). If they enter the long-term extension trial, they will not need to complete the follow-up visit in this trial. All other participants will return to the site for a follow-up visit (day 127 [±3 days]). For those participants who enter the long-term extension trial, adverse events and concomitant medication data will be recorded in the dose-ranging trial case report form (CRF) up until the date of extension trial randomization (defined as the trial completion date for the dose-ranging trial). For those participants who screen fail the long-term extension trial, adverse events and concomitant medication data will be recorded in the dose-ranging trial up until the follow-up visit (day 127 [±3 days]) CRF. Participants who complete the last scheduled visit will be considered to have completed the trial.

Participants who withdraw from the trial before completing the 14-week- treatment period will complete the assessments for the early termination visit and will be encouraged to complete the follow-up period in the trial.

The end of the trial is defined as the last visit of the last participant.

The total duration of participation in the trial is planned to be up to 24 weeks for each individual participant.

Trial procedures and assessments with respective time points are shown in Table 4 of the Trial Protocol.

2.2. Randomization and Blinding

This is a randomized, double-blind, placebo-controlled trial. Participants who meet all the inclusion criteria and none of the exclusion criteria will be randomly assigned to the 3 treatment arms: TEV-48574 900mg Q2W, TEV-48574 450mg Q2W or placebo in a 1:1:1 ratio, stratified by indication (UC or CD) and previous exposure to advanced therapy for IBD (yes/no) (biologics, JAK inhibitors, and S1P receptor modulators).

Table 1: Treatment Arms

Treatment Arm	Treatment (Q2W)
TEV-48574 900mg Q2W	TEV-48574 2250 mg (single loading dose)/900 mg (6 induction doses)
TEV-48574 450mg Q2W	TEV-48574 2250 mg (single loading dose)/450 mg (6 induction doses)
Placebo	Placebo

Q2W=every 2 weeks

Participants randomized prior to Amendment 03 with Revision 01 were randomized in a 1:1:1:1 ratio to TEV-48574 1800mg Q2W, TEV-48574 900mg Q2W, TEV-48574 450mg Q2W or placebo.

Participants will be randomly assigned to the treatment groups by means of a computer-generated randomization list using interactive-response technology.

Personnel who may be aware of IMP assignments will not be involved in the conduct of any trial procedures or assessment of any adverse events.

The participants, site, and sponsor trial team will be blinded until all participants complete the trial, and the database is locked for final analysis.

Pharmacokinetic, pharmacodynamics and immunogenicity samples will be collected for bioanalysis during the trial. Individuals responsible for sample bioanalysis and other responsible personnel may know who received test IMP and who received placebo IMP during the trial. Personnel responsible for bioanalysis will be provided with the randomization code to facilitate the analysis. However, the personnel responsible for bioanalysis will not have access to clinical safety and efficacy data and will provide concentration data to other personnel in a manner that will not identify individual participants (i.e., a dummy participant identifier will be linked to the concentration data of an individual participant).

The randomization list will be assigned to the relevant treatment groups through a qualified service provider, e.g., via the Randomization and Trial Supply Management (RTSM) system. The generation of the randomization list and management of the RTSM system will be done by a qualified service provider under the oversight of the responsible function at Teva. The specifications for randomization are under the responsibility and oversight of Teva Global Statistics.

Participant randomization codes will be securely maintained by the service provider contracted to generate the codes. At the time of interim and final analysis, after receiving unblinding request from Teva statistician, the service provider will provide the unblinded IMP assignment to the designated recipient(s) according to the processes defined in the relevant Standard Operating Procedure (SOP).

Interim analyses as described in Section 13 will be performed throughout the trial by unblinded analysis groups external to Teva and will be reviewed by an IDMC.

2.3. Data Monitoring Committee

An independent data review committee (IDMC) is established to monitor the trial while it is ongoing, including periodic reviews of safety data.

The IDMC charter provides details regarding conduct of safety reviews, including the pre-specified trial stopping rules for safety, and the procedures to protect the scientific integrity of trial.

Additional details regarding the safety IA statistical analysis are provided in Section 13.

2.4. Sequence of Planned Analyses

2.4.1. Planned Interim Analyses

Interim analyses for safety are planned for this trial (see Section 13). The interim analyses will be performed by an external, independent, unblinded reporting team.

The results will be reviewed by an IDMC comprised of external medical and statistical experts that will provide recommendations to the Sponsor. The IDMC charter provides details regarding procedures to protect the scientific integrity of the trial, conduct of the IA, and dissemination of results.

An interim efficacy/futility analysis was planned in each indication (UC and CD) when approximately 50% of patients in the indication have completed the 14-week primary efficacy readout time point or had the opportunity to do so. Due to accelerated recruitment, at the time that the data cut for this IA would have taken place, all participants in both indications had already been randomized. Given the duration of the treatment period in the trial, 14 weeks, Teva decided that the efficacy/futility IA would not be conducted as it would be impractical and have limited utility.

2.4.2. Final Analyses and Reporting

After the database has been locked for final analysis, all final analyses identified in this SAP will be performed.

The randomization codes will not be unblinded for final analysis until this SAP has been approved and issued and the database has been locked, including all external efficacy data (e.g. endoscopy scores) that are required for the primary and secondary efficacy analysis.

2.4.3. Analyses Out of the Scope of This SAP

Population Pharmacokinetic, Pharmacodynamic, and Pharmacokinetic/Pharmacodynamic Analysis: A pop-PK analysis will be performed, and, if feasible, a PK/PD analysis of relevant pharmacodynamic variables and/or exposure-response analysis of relevant efficacy or safety endpoints may be performed. The analyses may be reported in a separate pharmacometrics report and may have separate pre-specified analysis plan. These analyses may not be reported in the clinical study report (CSR) and are out of the scope of the current SAP.

Pharmacogenetic Analysis: Exploratory analyses may be undertaken to describe the relationship between genetic polymorphisms (single-nucleotide polymorphisms) and treatment effects of TEV-48574. These analyses may be reported in a separate report and may have a separate pre-specified analysis plan. These analyses may not be reported in the CSR and are out of the scope of the current SAP.

Biomarker Analysis: In addition to the analysis of hsCRP, fecal calprotectin, albumin, free and total TEV-48574 defined in this SAP, additional exploratory analyses may be undertaken to investigate the treatment effects of TEV-48574 on hsCRP, fecal calprotectin, autoantibodies, and other potential TL1A-mediated serum biomarkers. These additional analyses may be reported in a separate report and may have a separate pre-specified analysis plan. These analyses may not be reported in the CSR and are out of the scope of the current SAP.

2.5. Sample Size and Power Considerations

A total trial sample size of approximately 240 participants is planned, with 120 participants (40 participants per dose arm) each for the 2 indications of UC and CD.

The sample size does not include participants that were randomized to the 1800 mg Q2W treatment group prior to Protocol Amendment 03 with Revision 01.

Assumptions:

The clinical remission and endoscopic response rates assumed in the sample size and related operating characteristics computations are as follows:

- Clinical remission rates in UC group: 8% for placebo; 30% for TEV-48574 groups, respectively
- Endoscopic response rates in CD group: 12% for placebo; 34% for TEV-48574 groups, respectively

Additional scenarios for remission (UC), response rates (CD) are assumed as outlined in Table 2 to investigate the operating characteristics.

Table 2: Remission (UC), Response (CD) Rates Scenarios for Clinical Trial Simulations

Scenario	Indication	Placebo (%)	TEV-48574 450 mg Q2W (%)	TEV-48574 900 mg Q2W (%)
Null	UC	8	8	8
	CD	12	12	12

Table 2: Remission (UC), Response (CD) Rates Scenarios for Clinical Trial Simulations (Continued)

Scenario	Indication	Placebo (%)	TEV-48574 450 mg Q2W (%)	TEV-48574 900 mg Q2W (%)
Intermediate Efficacy	UC	8	20	20
	CD	12	25	25
Efficacy	UC	8	30	30
	CD	12	34	34
Efficacy-high dose only	UC	8	8	30
	CD	12	12	34

CD=Crohn's disease; Q2W=every 2 weeks; UC=ulcerative colitis.

Operating Characteristics:

Simulation-based operating characteristics for the analysis of the primary efficacy variable using a Bayesian Beta Binomial model are presented in Table 3. This model uses a posterior probability cutoff of ≥0.90 to declare success at the final analysis.

Table 3: Operating Characteristics of Final Analysis

	UC			CD		
	TEV-48574 450 mg Q2W	TEV-48574 900 mg Q2W	At least 1 dose	TEV-48574 450 mg Q2W	TEV-48574 900 mg Q2W	At least 1 dose
Null	7.9%	7.9%	13.5%	9.6%	9.6%	16.1%
Intermediate efficacy	61.1%	61.1%	77.3%	59.3%	59.3%	75.9%
Efficacy	91.0%	91.0%	97.6%	87.1%	87.3%	95.5%
Efficacy-high dose only	7.9%	91.1%	91.1%	9.6%	87.2%	87.3%

CD=Crohn's disease; Q2W=every 2 weeks; UC=ulcerative colitis.
Beta-Binomial model with non-informative prior Beta(1,1); decision rule for declaring difference from placebo is based on having a posterior probability of ≥0.90.
Note: Operating characteristics based on 500,000 simulations.

These simulations demonstrate that the probability of a false positive result, i.e., declaring success for at least 1 dose within each indication under the null assumption for both doses, is ≤16% for each indication, while the probability of declaring success for at least 1 dose within each indication under the efficacy assumption for both doses is ≥95%.

Since the analysis is done separately for each dose and each indication, the probability of demonstrating efficacy in one dose is not affected by the other dose. Moreover, no inflation of the Type 1 error rate is expected in hypothetical discordant scenarios where TEV-48574 is effective in one indication but not in the other.

As a supplemental analysis, a frequentist logistic regression analysis with 40 participants per dose per indication would provide at least 90% power using a Mantel-Haenszel test with a 1sided

nominal Type I error rate of 10% . However, the interpretation for this frequentist analysis is different from that of the Bayesian analysis as it is an inference on the sample space as opposed to the parameter space.

3. ANALYSIS SETS

3.1. Intent-to-Treat Analysis Set

The intent-to-treat (ITT) analysis set for each indication (UC or CD) will include all randomized participants.

In the ITT analysis set, participants will be categorized by treatment to which participants were randomized, regardless of which treatment they actually received.

Data collected from participants after treatment discontinuation will be included in the ITT analysis set.

3.2. Modified Intent-to-Treat Analysis Set

The modified intent-to-treat (mITT) analysis set for each indication (UC or CD) is a subset of the ITT analysis set including only participants who receive at least 1 dose of placebo, TEV-48574 450 mg Q2W or TEV-48574 900 mg Q2W.

In the mITT analysis set, participants will be categorized by treatment to which participants were randomized, regardless of which treatment they actually received.

3.3. Safety Analysis Set

The safety analysis set for each indication (UC or CD) will include all randomized participants who receive at least 1 dose of IMP.

In the safety analysis set, participants will be categorized by treatment actually received, regardless of the treatment to which they were randomized, unless otherwise specified.

Rules for assignment of treatment in case of mixed actual treatments:

- If a participant received placebo throughout the entire treatment period, then the actual treatment is placebo.
  - if a participant received a loading dose of TEV-48574 and received placebo in all other induction doses, the participant will be categorized to the TEV-48574 450 mg Q2W treatment arm.
- Otherwise (participant received at least 1 induction dose of TEV-48574), actual treatment is TEV-48574.
  - in case of kit mix-up, the participant will be categorized according to the highest induction dose (450 mg Q2W, 900 mg Q2W, or 1800 mg Q2W) that they received anytime during the trial

Rule for assignment of treatment in case of loading dose only (no induction doses):

- If a participant received a loading dose of TEV-48574 and had no induction doses, the participant will be categorized to the TEV-48574 450 mg Q2W treatment arm.

3.4. Immunogenicity Analysis Set

The immunogenicity analysis set will include all participants in the safety set who receive TEV-48574 and who have at least 1 reportable immunogenicity result.

3.5. Pharmacokinetic Analysis Set

The pharmacokinetic (PK) analysis set will include all participants in the safety set who receive TEV-48574 and who have at least 1 reportable serum concentration result.

Participants with kit mix-ups will be excluded from the PK analysis set.

4. GENERAL ISSUES FOR DATA ANALYSIS

4.1. General

Descriptive statistics for continuous variables include n, mean, standard deviation (SD), standard error (SE), median, minimum, and maximum. Descriptive statistics for categorical variables include participant counts and percentages, missing category will be displayed as appropriate.

In case of discrepancies between stratification factors as recorded in the clinical database and as recorded in the randomization system, the stratification factors as recorded in the clinical database will be used for all summaries and analyses.

4.2. Specification of Baseline Values

The baseline value is the last observed data before the first dose of trial drug, unless otherwise noted.

4.2.1. Baseline Value – MMS, CDAI, and PRO2

See Appendix 4 for derivation of baseline CDAI and PRO2-CD. The same window for participant-reported outcomes will be applied to baseline MMS and PRO2-UC.

Derived baseline MMS and derived baseline CDAI will not be used for eligibility or for definition of response for assignment to maintenance/re-induction in study 20038.

Derived baseline MMS and derived baseline CDAI will be used for sensitivity analysis only.

4.2.2. Baseline Value – ECG

The baseline value for electrocardiogram (ECG) variables is the mean of the triplicate values at the screening visit. The baseline value for the ECG finding (normal, abnormal) is the worst case finding of the triplicate findings.

4.3. Handling Withdrawals and Missing Data

Unless specified otherwise in this section, for all variables, only observed data will be used in the statistical analyses; i.e., there is no plan to estimate missing data.

All binary response and remission variables will be imputed as non-response if any of the following cases occur:

- Participants discontinuing treatment and/or trial during the 14-week treatment period regardless of the reason
- Participants who used rescue medications for UC or CD prior to their week 14 evaluation of remission/response (see Section 4.5 for classification rules)
- Participants undergoing major UC or CD related surgery (see Section 4.6 for classification rules)

4.3.1. Missing Week 14 Endoscopy Assessment

All binary response and remission variables that are based on the week 14 endoscopy will be imputed as non-responder/non-remission if the week 14 endoscopy assessment is not available for the participant. In particular,

- UC participants who did not complete evaluation for clinical remission at week 14 will be reported as non-remitters.
- CD participants who did not complete evaluation for endoscopic response at week 14 will be reported as non-responders.

4.3.2. Missing items – MMS, CDAI, PRO2-UC, and PRO2-CD

The MMS, CDAI, PRO2-UC and PRO2-CD are composite scores (see derivation rules in Appendix 4, Appendix 5). The patient-reported items of stool frequency (UC and CD), rectal bleeding (UC), and abdominal pain (CD) are based on e-diary data collected in a 7-day window before the assessment date of the score, which may be extended to 10 days prior to the assessment date, as described in Appendix 4. A minimum of 3 consecutive days of completed diary entries or 4 nonconsecutive days are necessary for each parameter, otherwise the corresponding item will be considered missing.

If any item of a composite score is missing, then the score will be considered missing.

If a score is missing, then the participant will be considered a non-responder /non-remitter for all binary outcomes based on this score, except for intermittent missing data visits that will be excluded from that visit analyses.

4.3.3. Missing Items – Robarts Histopathology Index

If any of the 4 subscores are missing, then the final score is set to missing. If the final score is missing, then the participant's response will be imputed as a non-remitter for week 14.

4.3.4. Missing Items – Geboes Index

If any of the items are missing, then the score is set to missing. If the score is missing, then the participant's response will be imputed as a non-remitter for week 14.

4.3.5. Missing Items – Global Histologic Activity Score

If any of the 8 items are missing, then the total score is set to missing. If the total score is missing, then the participant's response will be imputed as a non-responder for week 14.

4.4. Trial Days and Visits

For by-visit summaries, if there are multiple assessments at a postbaseline visit then the last non-missing assessment at that visit will be used for the summary (this includes scheduled and unscheduled assessments).

Trial days are numbered relative to the first day of trial drug administration. The start of treatment (Day 1) is defined as the date on which a participant takes the first dose of trial drug, as recorded on the clinical database. Days will be numbered relative to treatment start (i.e., ..., –2, –

1, 1, 2, ...; with day 1 being the first day of trial drug administration and day –1 being the day before the first day of trial drug administration).

4.5. Classification of Intercurrent Events – Initiation or Dose Increase of Inflammatory Bowel Disease Medications

The intercurrent events of initiation or dose increase medications used to treat the IBD will be classified as follows:

Medication	ICE Classification
5-ASA; sulfasalazine (SSZ)	Important protocol deviations, including use or changes in other medications (treatment policy)
Immunosuppressants – other than MTX	Important protocol deviations, including use or changes in other medications (treatment policy)
Immunosuppressants – MTX	Rescue medication
Systemic corticosteroids (IV, PO, IM, suppository)	Rescue medication
Advanced IBD therapies (biologics, JAK inhibitors, S1P inhibitors)	Rescue medication

As described in the primary estimand (Section 1.2), ICEs classified as "Rescue medication" will be handled using the composite strategy, i.e., the participant will be reported as non-remitter/non-responder; ICEs classified as "Important protocol deviations, including use or changes in other medications" will be handled using the treatment policy, i.e., data after the ICE will be used regardless of the ICE. These strategies will be applied to all response/remission efficacy endpoints.

All participants with initiation or dose increase of IBD medications will be reviewed in the statistical data review meeting prior to DBL, where the classification will be confirmed on a case-by-case basis.

4.6. Classification of Intercurrent Events – Inflammatory Bowel Disease Related Surgery

All IBD-related Major surgery will be considered as a major UC/CD related surgery that will be handled using composite strategy. Classification of surgeries will be done on a case-by-case basis and reviewed in the blinded statistical data review meeting prior to DBL.

As described in the primary estimand (Section 1.2), ICEs classified as "Major UC/CD related surgery" will be handled using the composite strategy, i.e., the participant will be reported as non-remitter/non-responder. This strategy will be applied to all response/remission efficacy endpoints.

Note that participants undergoing IBD surgery will likely require perioperative systemic corticosteroids and thus will be treated using the composite strategy, i.e., these participants will be reported as non-remitters/non-responders, on that basis.

5. TRIAL POPULATION

5.1. General

The ITT analysis set will be used for all trial population summaries. Summaries will be presented by treatment group, pooled TEV-48574, and all participants, separately by indication (UC or CD) and pooled (UC and CD).

5.2. Participant Disposition

Participants screened, participants screened but not randomized, will be summarized only for the total group (all participants) using participant counts.

Participants in the ITT, mITT, safety, and immunogenicity analysis set, participants who complete the 14-week treatment period, and participants who complete the trial and rollover into the LTE trial or complete the trial via the follow-up period will be summarized using descriptive statistics. The denominator for calculating the percentages will be the number of participants in the ITT analysis set.

Participants who withdraw from the trial prior to completion of the 14-week treatment period and participants who withdraw from the trial during the follow-up period after completing the treatment period, if applicable, will also be summarized using descriptive statistics by reason for withdrawal.

5.3. Demographics and Baseline Characteristics

The continuous variables of participant age, weight, height, and body mass index (BMI), and will be summarized using descriptive statistics. The categorical variables of participant sex, race, ethnicity, BMI category (18.5<; ≥18.5 and <25; ≥25 and <30; ≥30), smoking status (current or former smoker yes/no), previous exposure to advanced therapy for IBD (yes/no), and region (North America/Eastern EU - Poland/Eastern EU – Other/Western EU/Japan/Israel) will be summarized using descriptive statistics for each category. Missing categories will be presented if necessary.

In addition, a summary of participants by region and country will be provided by indication and overall.

5.4. Disease Characteristics

The following disease characteristics variables at baseline will be summarized using descriptive statistics: time since initial IBD diagnosis (years), MMS and components (UC), PRO2-UC score (UC), SES-CD (CD), CDAI score, and stool frequency and abdominal pain components (CD), PR02-CD (CD), fecal calprotectin value, hsCRP value, and albumin value.

5.5. Baseline IBD Treatments

The incidence of baseline concomitant IBD therapies will be summarized using descriptive statistics using the following categories:

  - Glucocorticosteroids
  - 5-ASA
  - Immunosuppressants

UC and CD treatment categories definitions will be reviewed in the blinded data review meeting prior to DBL, and additional categories/medications may be added.

In addition, baseline use of corticosteroids and dose will be summarized using descriptive statistics for equivalent prednisone dose, as follows:

- Descriptive statistics of the dose (min, max, mean, etc.)
- Incidence of current treatment with oral corticosteroids (Yes/No), and a breakdown to dose categories:
  - <10mg
  - 10-20mg
  - >20mg
  - Budesonide
    - 3mg to < 6 mg
    - 6 mg to ≤ 9 mg
    - > 9 mg

The conversion table of oral corticosteroid doses to equivalent prednisone dose is provided in Appendix 6.

5.6. Medical History

All medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The incidence of medical history abnormalities will be summarized using descriptive statistics by system organ class (SOC) and preferred term (PT). Participants are counted only once in each SOC category, and only once in each PT category.

5.7. Inflammatory Bowel Disease History

The continuous variable of number of hospitalizations due to UC or CD will be summarized using descriptive statistics. The categorical variables of histology findings to support the diagnosis of UC/CD (yes/no), colonic stricture (yes/no), and history of colonic or small bowel obstruction (yes/no) will be summarized using descriptive statistics for each category. Missing categories will be presented if necessary.

5.8. Prior Therapy and Medication

Any prior therapy, medication, or procedure for IBD (anytime in the past) will be recorded in the case report form (CRF). Any other prior therapy, medication, or procedure a participant has had within 4 weeks prior to the screening visit will be recorded in the CRF. Generic or trade name,

indication, and dosage will be recorded. The sponsor will encode all therapy and medication according to the World Health Organization (WHO) drug dictionary (WHO Drug).

Prior therapies and medications will include all medications taken and therapies administered before the first day of trial drug administration.

The incidence of prior therapies and medications will be summarized using descriptive statistics by therapeutic class and PT. Participants are counted only once in each therapeutic class category, and only once in each PT category.

In addition, the incidence of prior therapies, medications, or procedures for IBD will be summarized using descriptive statistics using the following categories:

- Glucocorticosteroids
- 5-ASA
- Immunosuppressants
- TNF inhibitors
- Integrin inhibitors
- IL-12/23 inhibitors
- Anti-IL-23
- JAK inhibitors
- S1P Modulators
- Investigational drugs

UC and CD treatment categories will be reviewed in the blinded data review meeting prior to DBL, and additional categories/medications may be added. Investigational drugs will be classified on a case-by-case basis.

The number of advanced therapies for IBD that a participant was exposed to in the past will be summarized using descriptive statistics.

5.9. Electrocardiography

ECG numeric parameters (intervals) and interpretation (normal, abnormal not clinically significant, and abnormal clinically significant) at screening will be summarized using descriptive statistics.

5.10. Trial Protocol Deviations

Data from participants with any important protocol deviations during the trial will be summarized overall and for each category using descriptive statistics.

5.11. Intercurrent Events

The incidence of the intercurrent events for which the composite strategy is used in the primary analysis (see Section 1.2 for primary estimand and Sections 4.5 and 4.6 for classification rules) will be summarized using descriptive statistics. The summary will be presented for the mITT analysis set.

6. EFFICACY ANALYSIS

6.1. General

The mITT analysis set will be used for all efficacy analyses. Summaries will be presented by treatment group and indication (UC or CD).

Descriptive statistics for all efficacy endpoints will be summarized for the mITT analysis set. In addition, descriptive statistics will be provided for the ITT analysis set, which includes participants in the 1800 mg treatment group that were randomized to the trial prior to PA03 Rev01.

6.1.1. Efficacy Endpoints of Response and Remission at Week 14

The efficacy endpoints of response and remission at week 14 are presented for UC and CD in Table 4 and Table 5.

Table 4: Response and Remission Endpoints at Week 14 – UC

Endpoint	Definition	Classification
Clinical remission (MMS)	MMS of ≤2 points defined by: a stool frequency subscore of 0 or 1, rectal bleeding subscore of 0, an endoscopic subscore of 0 or 1, where a score of 1 does not include "friability"	Primary
Clinical remission 2 (MMS)	MMS of ≤2 points defined by: a stool frequency subscore of 0 or 1 and not greater than the baseline, rectal bleeding subscore of 0, an endoscopic subscore of 0 or 1, where a score of 1 does not include "friability"	Supplementary to primary
Clinical response (MMS)	≥ 2 points decrease in MSS AND at least a 30% reduction from baseline with either a decrease in rectal bleeding subscore of at least 1 or an absolute rectal bleeding subscore of less than or equal to 1	Secondary
Endoscopic improvement (Mayo endoscopic subscore)	Mayo endoscopic subscore of 0 or 1, where a score of 1 does not include "friability"	Secondary
Endoscopic remission (Mayo endoscopic subscore)	Mayo endoscopic sub-score of 0	Secondary
Clinical response (PRO2)	≥ 50% decrease in PRO2 from baseline	Secondary

Table 4: Response and Remission Endpoints at Week 14 – UC (Continued)

Endpoint	Definition	Classification
Clinical remission (PRO2)	rectal bleeding=0, and stool frequency=0	Secondary
Histological remission (RHI)	RHI ≤5	Exploratory
Histological remission (Geboes Index)	Geboes index score ≤3.1	Exploratory
Histologic-Endoscopic Mucosal Improvement (HEMI)	Mayo endoscopic subscore of 0 or 1 (where a score of 1 does not include "friability") and Geboes score ≤3.1	Exploratory

Table 5: Response and Remission Endpoints at Week 14 – CD

Endpoint	Definition	Classification
Endoscopic response (SES-CD)	≥ 50% decrease in SES-CD from baseline	Primary
Endoscopic response 2 (SES-CD)	> 50% decrease in SES-CD from baseline, or a decrease of at least 2 points for subjects with a baseline score of 4 and isolated ileal disease	Supplementary to primary
Endoscopic response (MM-SES-CD)	≥ 50% decrease in MM-SES-CD from baseline	Secondary
Clinical response (CDAI)	≥100-point decrease in CDAI score from baseline	Secondary
Clinical remission (CDAI)	CDAI score <150	Secondary
Clinical response (PRO2)	≥ 50% decrease in PRO2 from baseline	Secondary
Clinical response 2 (PRO2)	≥ 30% decrease in PRO2 from baseline	Exploratory
Clinical remission (PRO2)	abdominal pain ≤1, and stool frequency ≤3	Secondary
Clinical remission 2 (PRO2)	abdominal pain ≤1, and stool frequency ≤1.5	Exploratory
Clinical remission 3 (PRO2)	abdominal pain ≤1, and stool frequency ≤2.8	Exploratory
Endoscopic remission (SES-CD)	SES-CD score of 0-2	Exploratory
Endoscopic remission 2 (SES-CD)	SES-CD score of 0-4 with no individual subscore >1	Exploratory
Histologic response (GHA)	≥50% decrease in GHA Score from baseline	Exploratory

Similar secondary/exploratory response and remission endpoints based on PRO2 and CDAI are defined at all timepoints where they are assessed. Details regarding the CDAI and PRO2-CD scores are provided in Appendix 4 and Appendix 5.

6.2. Primary Efficacy Endpoints and Analysis

6.2.1. Primary Endpoint

The primary endpoint is response (yes or no) at week 14, where response in UC participants is defined as clinical remission (based on MMS) and response in CD participants is defined as endoscopic response (based on SES-CD).

6.2.1.1. Primary Endpoint in Ulcerative Colitis – Clinical Remission

The primary efficacy endpoint in UC participants is clinical remission at week 14 based on the modified Mayo score (MMS), defined as follows: MMS ≤ 2 with

- stool frequency subscore of 0 or 1, and
- rectal bleeding subscore of 0, and
- endoscopic subscore of 0 or 1, where a score of 1 does not include "friability"

Details regarding the endoscopy and the MMS are provided in Section 6 of the trial protocol.

6.2.1.2. Primary Endpoint in Crohn's Disease – Endoscopic Response

The primary efficacy endpoint in CD participants in endoscopic response at week 14, defined as a reduction from baseline in Simple Endoscopic Score for Crohn's Disease (SES-CD) of at least 50%.

6.2.2. Primary Efficacy Analysis

A Beta-binomial model with a non-informative Beta (1,1) prior will be used to analyze the primary response endpoint, clinical remission (for UC) or endoscopic response (for CD), hereafter denoted as "response". The model will be fit separately for each dose within each indication.

Inference will be based on posterior distribution of the response rates. A TEV-48574 dose will be declared successful at the final analysis if the posterior probability that the response rate in the TEV-48574 dose is higher than the response rate in the placebo arm is ≥ 0.90. Put another way, the success criterion for each dose within indication can be written as:

Posterior Probability (TEV-48574 response rate – placebo response rate >0) ≥ 0.90

The posterior distribution of (TEV-48574 response rate – placebo response rate) for each treatment group will be presented graphically. In addition, the following quantities will be presented:

- Posterior means of the response rates for TEV-48574 900 mg Q2W, TEV-48574 450 mg Q2W and placebo
- Two-sided 95% credible intervals for difference in response rates (TEV-48574 doses – placebo)

- For each TEV-48574 dose, Posterior Probability (TEV-48574 response rate – placebo response rate > 0)

Appendix 1 provides details on the Beta-binomial model.

6.2.3. Sensitivity and Supplementary Analysis

The following sensitivity and supplementary analyses of the primary endpoint will be performed:

1. Tipping point analysis will be performed to test the robustness of the primary analysis conclusion to the non-response imputation of participants that discontinue the trial or have missing week 14 MMS (UC participant) or SES-CD (CD participants) data used in the primary analysis. The tipping point analysis will be conducted by varying the number of participants imputed as non-responders in the active and placebo groups.
2. A supplementary analysis for the primary endpoint using the response variables defined in the first supplementary estimand, clinical remission 2 (MMS) for UC and endoscopic response2 (SES-CD) for CD (see Table 5), will be performed. The primary analysis model will be used in this analysis.
3. A supplementary analysis for the primary efficacy variable will be performed using a logistic regression model fit separately to each indication with the following fixed effects: dose (as categorical variable), actual previous exposure to advanced IBD therapy (yes/no), and baseline MMS (UC participants) or SES-CD (CD participants).
   - Each comparison of TEV-48574 dose versus placebo will be tested (one-sided-) at a nominal significance level of α=0.1.
   - Odds ratios (for each of the TEV-48574 doses versus placebo), associated two-sided 95% confidence intervals (CIs), and nominal one-sided- p-values will be reported.
4. Cochran-Mantel-Haenszel (CMH) test for pairwise comparisons between the active dose groups and placebo separately for UC and CD. Odds ratios (for each of the TEV-48574 doses versus placebo), difference in response rate, and the associated two-sided 95% CIs and nominal one-sided- p-values will be reported.
5. Cochran-Mantel-Haenszel (CMH) test for pairwise comparisons between the active dose groups and placebo separately for UC and CD, stratified by actual previous exposure to advanced IBD therapy (yes/no). The model will be used if data permits. Odds ratios (for each of the TEV-48574 doses versus placebo), difference in response rates, and the associated two-sided 95% CIs and nominal one-sided- p-values will be reported.
6. Bayesian logistic regression model for both UC and CD using a non-informative prior with a logistic regression model fit separately to each indication with the following fixed effects: dose (as categorical variable) and actual previous exposure to advanced IBD therapy (yes/no). The model will utilize a Jeffrey's prior which is a beta (1/2, 1/2).

6.2.4. Subgroup Analysis

The primary endpoints (UC and CD) will be evaluated using a Beta-binomial model with a non-informative Beta (1,1) prior as described in Section 6.2.2 for the subgroups below. The two-sided 95% credible intervals will be provided. A forest plot of the 95% two-sided credible intervals will be presented.

The primary endpoints (UC and CD) will also be evaluated for the subgroups using a logistic regression model as described in Section 6.2.3. Odds ratios (for each of the TEV-48574 doses versus placebo) and associated two-sided 95% CIs will be reported. A forest plot of the two-sided 95% CIs will be presented.

If there are <20 participants per subgroup within each indication, only descriptive statistics will be provided.

- Previous use of advanced IBD therapies (yes/no)
- Region (North America/Eastern EU - Poland/Eastern EU – Other/Western EU/Japan/Israel)
- Sex (female/male)
- BMI category (18.5<; ≥18.5 to <25; ≥25 to <30; ≥30)
- Weight (<median, ≥median)
- Age category (18-40; 41-64; >65)
- Smoking status (current or former smoker: yes/no)
- Use of corticosteroids at baseline (yes/no)

Note: the median weight will be calculated for each indication separately.

For the subgroup analysis by previous use of advanced IBD therapies (yes/no), the posterior distributions of (TEV-48574 response rate – placebo response rate) in each stratum will be presented graphically for each treatment group. The logistic regression model will not include previous use of advanced IBD therapies (yes/no) as a covariate.

6.3. Secondary Efficacy Endpoints and Analysis

6.3.1. Secondary Endpoints

The secondary efficacy endpoints are presented in Section 1.1. The response and remission secondary endpoints at Week 14 are presented in Table 4 and Table 5.

6.3.1.1. Secondary Endpoints Analysis

All secondary efficacy endpoints will be analyzed separately for UC and CD participants.

All binary secondary efficacy endpoints (response/remission/improvement at week 14 and clinical response based on CDAI at weeks 4, 8, and 12) will be analyzed using the following methods:

1. Beta-Binomial model with non-informative Beta (1,1) prior, as described in Section 6.2.2.
2. Logistic regression model as described in Section 6.2.3.
3. Cochran-Mantel-Haenszel (CMH) test for pairwise comparisons between the active dose groups and placebo separately for UC and CD, as described in Section 6.2.3.
4. Cochran-Mantel-Haenszel (CMH) test for pairwise comparisons between the active dose groups and placebo separately for UC and CD, stratified by actual previous exposure to advanced IBD therapy (yes/no), as described in Section 6.2.3. The model will be used if data permits.

For the Bayesian analysis, the posterior means of the response rates, two-sided 95% credible intervals for difference in response rates, and the posterior probability that (TEV-48574 response rate – placebo response rate > 0) will be reported.

For the frequentist analyses (logistic regression and CMH), the odds ratios, difference in response rates, and the associated two-sided 95% confidence intervals (CIs) and nominal one-sided- p-values will be reported.

6.4. Exploratory Efficacy Endpoints and Analysis

The exploratory efficacy endpoints are presented in Section 1.3. The response and remission exploratory endpoints at Week 14 are presented in Table 4 and Table 5.

Binary endpoints will be summarized using descriptive statistics for categorical data. In addition, a graph of response rate at each visit will be presented where applicable.

Summary statistics for PRO2-UC items, PRO2-UC total score, PRO2-CD items, PRO2-CD total score, and CDAI score will be presented by visit. Actual values and changes from baseline to each visit will be provided. In addition, graphs of mean change from baseline to each visit by treatment group will be presented.

6.4.1. Efficacy Biomarkers

The following efficacy biomarkers endpoints will be summarized as part of the statistical analysis of exploratory efficacy endpoints:

- hsCRP:
  - hsCRP change from baseline at each visit
  - hsCRP percent change from baseline at each visit
  - hsCRP <5 mg/L, hsCRP ≥5 mg/L at each visit
- fecal calprotectin:
  - fecal calprotectin change from baseline at each visit
  - fecal calprotectin percent change from baseline at each visit
  - fecal calprotectin ≤250 mg/kg, fecal calprotectin >250 mg/kg at each visit

- albumin:
  - albumin change from baseline at each visit
  - Albumin percent change from baseline at each visit
  - albumin ≤35 g/L, albumin >35 g/L at each visit

Summary statistics for hsCRP fecal calprotectin, and albumin will be presented by visit. Actual values, change from baseline, and percent change from baseline to each visit will be provided. In addition, graphs of median value, median change from baseline and median percent change from baseline to each visit by treatment group will be presented. Box plots of hsCRP, fecal calprotectin, and albumin will be presented.

The analysis will be repeated by response at week 14 (yes/no) as defined by the primary endpoint.

7. MULTIPLE COMPARISONS AND MULTIPLICITY

No adjustments to control the false-positive rate will be made for the preplanned multiple comparisons/endpoints analyses in this phase II study.

8. SAFETY ANALYSIS

8.1. General

The safety analysis set will be used for all safety analyses. Summaries will be presented by treatment and pooled TEV-45874, separately by indication (UC or CD) and pooled (UC and CD). The summaries will include participants in the 1800 mg treatment group that were randomized prior to PA03Rev01.

If an indication is stopped for safety reasons (see Section 13.1), pooled analyses of UC and CD will not be presented.

Safety listings will include all data collected, starting from screening. Safety summary tables will include all treatment-emergent data, ie, data collected from first day of trial drug until the end of the trial (including the FU period), unless specified otherwise.

8.2. Duration of Exposure to Trial and Trial Drug

Duration of trial (days in trial) is the number of days starting from Day 1 to the last visit day in the trial (last visit day – Day 1 + 1).

Duration of exposure is the number of days starting from the 1st day of trial drug to 4 weeks after last dosing (4 weeks after last dosing – first day of trial drug + 1).

Number of doses received (planned dose schedule is 1 loading dose and 6 induction doses) will be summarized as categorical data using descriptive statistics.

Weeks on treatment will be summarized as categorical data using the categories ≥2 weeks, ≥4 weeks, ≥6 weeks, ≥8 weeks, ≥10 weeks, and ≥12 weeks using descriptive statistics. Duration of treatment (days) will also be summarized as continuous data using descriptive statistics.

8.3. Adverse Events

All adverse events (AEs) will be coded using MedDRA. Common Terminology Criteria for Adverse Events (CTCAE) will be used to grade the severity of AEs. Treatment-emergent adverse events (TEAEs) are defined as AEs that start either on or after the first dose of the study drug and until first dose in study TV48574-IMM-20038 (if enrolled) or end of the FU period of the current study.

An overview table summarizing the following AE categories will be presented:

- Participants with at least 1 TEAE
- Participants with at least 1 TEAE CTCAE grade 3/4
- Participants with at least 1 treatment-related TEAE (related to IMP)
- Participants with at least 1 device-related TEAE
- Participants with at least 1 treatment-related TEAE CTCAE grade 3/4
- Participants with at least 1 serious TEAE
- Participants with at least 1 serious IMP related TEAE

- Participants with at least 1 serious device related TEAE
- Participants with at least 1 TEAE leading to discontinuation from the trial
- Participants with at least 1 TEAE leading to dose interruption
- Participants with at least 1 TEAE leading to treatment discontinuation
- Participants who died
- Participants with at least 1 protocol-defined TEAE of special interest (AESI)

Summaries will be presented for all TEAEs, and for the TEAE categories listed above.

The incidence of TEAEs will be summarized (overall and by grade) using descriptive statistics by SOC and PT (all TEAEs overall will also be presented by PT category in descending order of incidence in the pooled TEV-48574 group). Participants are counted only once in each SOC category, and only once in each PT category. For the summaries by grade, participants are counted at the highest grade. TEAEs that are missing information on seriousness will be excluded from the summary of serious TEAEs but included in the summary of non-serious TEAEs.

A Beta-binomial analysis of the 4 types of AEs used to define the stopping criteria in the interim safety analyses (see Section 13.1 and Appendix 1) will be conducted using the final data. The posterior probability that the AE rate in each active group and the 3 dose levels combined is larger than the AE rate in the placebo group will be reported for UC and CD separately and across the entire trial population.

Participant listings of all AEs, serious AEs, serious IMP related AEs, serious device related AEs, adverse events leading to discontinuation from the trial, and protocol-defined AESIs will be presented.

If any participant dies during the trial, a listing of deaths will be provided, and all relevant information will be discussed in the participant narrative included in the CSR.

8.4. Clinical Laboratory Tests

Laboratory test results will be presented in standard international (SI) units.

Summary statistics for chemistry, hematology, and urinalysis laboratory tests will be presented at baseline and each visit that this is measured. Laboratory values and changes from baseline to each visit will be summarized using descriptive statistics. Listings of all individual participants' laboratory tests will be presented.

Shifts (below, within, and above the normal range) from baseline to each visit that this is measured will be summarized using participant counts.

Summaries of potentially clinically significant abnormal values will include all postbaseline values (including scheduled, unscheduled, and early termination visits). The incidence of potentially clinically significant abnormal values will be summarized for laboratory variables using descriptive statistics with the criteria specified in Table 6.

Table 6: Criteria for Potentially Clinically Significant Laboratory Values

Test Category	Test	Criterion value
Serum chemistry	ALT	>5 × ULN >5 × BL if BL is abnormal (> ULN)
	AST	>5 × ULN >5 × BL if BL is abnormal (> ULN)
	ALP	>5 × ULN >5 × BL if BL is abnormal (> ULN)
	Creatinine	>3 ×ULN >3 × BL if BL is abnormal (> ULN)
	Bilirubin (total)	>3 × ULN >3 × BL if BL is abnormal
	Potassium	<3.0 mmol/L >6.0 mmol/L
	Sodium	<120 mmol/L >155 mmol/L
	Calcium	Corrected serum calcium of >12.5 mg/dL Corrected serum calcium of <7.0 mg/dL
	Phosphate	<1.4 mg/dL
Hematology	Hemoglobin	<8.0 g/dL and/or decrease of 3.0 g/dL from BL > ULN + 4.0 g/dL
	WBC counts	<2.0 x 10^9/L >100 x 10^9/L
	Neutrophil	<0.5 x 10^9/L
	Lymphocyte	<0.5 x 10^9/L
	Platelet counts	<50 x 10^9/L
Urinalysis	HGB	≥2 unit increase from BL
	Glucose	≥2 unit increase from BL
	Ketones	≥2 unit increase from BL
	Total protein	≥2 unit increase from BL

ALP=alkaline phosphatase; ALT=alanine aminotransferase; ANC=absolute neutrophil count AST=aspartate aminotransferase; BL = baseline; BUN=blood urea nitrogen; GGT=gamma-glutamyl transpeptidase; HGB=hemoglobin; LDH=lactate dehydrogenase; ULN=upper limit of normal range; WBC=white blood cell.

8.4.1. Liver-Related Laboratory Tests

Maximum post-baseline elevation in ALT, AST and bilirubin will be summarized as incidence and shift tables using the criteria specified in Table 7.

Table 7: Liver-Related Laboratory Tests Criteria

Test	Criterion Value
ALT	>3 x ULN
	>5 x ULN
	>10 x ULN
	>20 x ULN
AST	>3 x ULN
	>5 x ULN
	>10 x ULN
	>20 x ULN
ALT or AST	>3 x ULN
	>5 x ULN
	>10 x ULN
	>20 x ULN
Bilirubin (total)	≥2 x ULN
Hy's Law	{ALT ≥3 x ULN or AST ≥3 x ULN} and {bilirubin (total) ≥2 x ULN} and {ALP <2 x ULN}

Incidence of participants meeting laboratory criteria for Hy's Law, i.e. {ALT ≥3 x ULN or AST ≥3 x ULN} and {bilirubin (total) ≥2 x ULN} and {ALP <2 x ULN}, will be listed.

8.4.2. Other Clinical Laboratory Tests

Other clinical laboratory tests results will be presented in data listings.

8.5. Vital Signs and Weight

Summary statistics for pulse, systolic and diastolic blood pressure, respiratory rate, body temperature, and weight will be presented at baseline and each visit that this is measured. Vital signs values and changes from baseline to each visit will be summarized using descriptive statistics.

Summaries of potentially clinically significant abnormal values will include all postbaseline values (including scheduled, unscheduled, and early termination visits). The incidence of potentially clinically significant abnormal values will be summarized using descriptive statistics with the criteria specified in Table 8.

Table 8 specifies the criteria for identifying vital signs as potentially clinically significant abnormal values. Note that to qualify as potentially clinically significant abnormal, a value needs to meet both criteria below, i.e., have a value beyond the criterion value and a change of at least the magnitude specified in the change relative to baseline column.

Table 8: Criteria for Potentially Clinically Significant Vital Signs

Vital Sign	Criterion value	Change relative to baseline
Pulse	≥120 bpm	Increase of ≥15 bpm
	≤50 bpm	Decrease of ≥15 bpm
Systolic blood pressure	≥180 mm Hg	Increase of ≥20 mm Hg
	≤90 mm Hg	Decrease of ≥20 mm Hg
Diastolic blood pressure	≥105 mm Hg	Increase of ≥15 mm Hg
	≤50 mm Hg	Decrease of ≥15 mm Hg
Respiratory rate	<10 breaths/min	
Body temperature	≥38.3°C	Change of ≥1.1°C

bpm=beats per minute

Maximum post-baseline changes in weight will be summarized using the criteria specified in Table 9. Exclusion of participants with a baseline weight ≥ 100kg and/or BMI ≥30 at baseline from this analysis will be discussed in the blinded data review meeting prior to DBL and unblinding.

Table 9: Criteria for Weight Increase or Decrease

Test	Criterion Value
Weight	Increase of ≥5% from BL
	Increase of ≥10% from BL
	Decrease of ≥5% from BL
	Decrease of ≥10% from BL

BL = baseline

8.6. Electrocardiography

Shifts (normal, abnormal not clinically significant, and abnormal clinically significant) from baseline to each visit that this is measured, and from baseline to overall (anytime) will be summarized using participant counts for the investigator interpretation. For overall result interpretation the worst postbaseline finding for the participant (the abnormal finding if there are both normal and abnormal findings) will be used in the summaries.

Summary statistics for ECG variables will be presented at baseline and each visit that it is measured. Actual values and changes from baseline to each visit will be summarized using descriptive statistics.

The incidence of potentially clinically significant abnormal values will be summarized using descriptive statistics with the criteria specified in Table 10. Note that to qualify as potentially clinically significant abnormal, a value needs to meet both criteria below, if applicable: e.g., have a value beyond the criterion value and a change of at least the magnitude specified in the change relative to baseline column. A listing of potentially clinically significant ECG results will be presented.

Table 10: Criteria for Potentially Clinically Significant Abnormal ECG

ECG Test	Criterion Value	Change Relative to Baseline
PR Interval	≥200 msec	Increase of at least 25% from baseline
QRS Duration	≥110 msec	Increase of at least 25% from baseline
QTcF Interval	≥450 msec and <500 – male or ≥460 msec and <500 – female	≥30 and <60 msec increase from baseline
	≥450 msec and <500 – male or ≥460 msec and <500 – female	≥60 msec increase from baseline
	≥500 msec	≥30 and < 60 msec increase from baseline
	≥500 msec	≥60 msec increase from baseline

8.7. Concomitant Medications or Therapies

Concomitant therapies and medications, including medications that are taken on an as needed basis and occasional therapies, will be monitored during the trial. Details of prohibited medications may be found in Section 5.6.4 of the trial protocol. All concomitant medications will be coded using the WHO Drug dictionary.

Any medication or therapies initiated on the same date of study treatment exposure will not be considered as concomitant medications or therapies. This also applies for the same year or month if other specific information is missing. In addition, any medications or therapies that are recorded but have completely missing start and stop dates will not be considered concomitant. A separate listing will be provided to account for these medications or therapies recorded in the study, but not considered concomitant.

The incidence of concomitant therapies and medications will be summarized using descriptive statistics by therapeutic class category and PT. Participants are counted only once in each therapeutic class, and only once in each PT category. Concomitant therapies and medications will include all medications taken during the treatment period. Any medications or therapies started after the treatment period will be included in a separate listing.

For patients who early terminate, any medication that started prior to the ET visit or up to 2 weeks after last dose of IMP, whichever comes first, is considered concomitant. Any medication that started at or after ET visit is not considered concomitant and will not be included in the summary table but will be included in the listing and footnoted accordingly.

8.8. Other Safety Assessments

8.8.1. Local Tolerability

Local tolerability at the injection site (erythema, ecchymosis, induration, tenderness, warmth, and swelling) will be assessed using standardized scales as described in Table 8 of the trial protocol.

Local tolerability response will be summarized at each visit that this is measured using descriptive statistics as categorical data.

8.8.2. Injection Pain Intensity Assessment

Injection site pain will be reported using a standardized 11-point pain intensity numerical response scale (NRS-11) where 0 is “No pain” and 10 is the “Worse possible pain”

The injection site pain intensity assessment will be summarized at each timepoint and visit that it is measured using descriptive statistics as categorical data. The timepoints during IMP administration and ‘after the completion of IMP administration’ will be presented as one timepoint, based on the worst assessment.

9. TOLERABILITY VARIABLES AND ANALYSIS

Tolerability will be assessed by the number (%) of participants who withdraw from the trial due to adverse events. Tolerability is summarized with participant disposition.

If more than 10% of participants withdraw from the trial due to adverse events, a Kaplan-Meyer analysis of time to withdrawal will be performed, separately by indication (UC or CD) and pooled (UC and CD).

In case that any dose is stopped for safety reasons (see Section 13.1), the tolerability analysis will be repeated excluding the dose(s) that was (were) stopped.

10. IMMUNOGENICITY ANALYSIS

Results of immunogenicity assessments will be listed and summarized using descriptive statistics. Immunogenicity summaries will be presented by indication and treatment group, including pooled TEV-48574.

Participants with pre-existing ADAs correspond to participants with ADAs present at baseline (prior to first dose of IMP). Participants with missing ADA sample at baseline will be considered as without pre-existing ADA.

ADA status at each post-baseline visit will be classified to one of 3 categories:

- Treatment-emergent ADA positive:
  - the participant has a positive ADA sample (after 1st dose of IMP) but not at baseline (prior to first dose of IMP), or
  - the participant had a positive ADA sample at baseline (prior to first dose of IMP) and at the visit (after 1st dose of IMP), with at least a 4-fold increase in titer level.
- ADA negative: the participant does not have a positive ADA sample at the visit
- Inconclusive: otherwise

ADA status of each participant (anytime post-baseline) will be classified as follows:

- Participants with treatment-emergent ADA correspond to participants with at least one treatment-emergent positive ADS sample.
- Participants with unclassified ADA correspond to patients with pre-existing ADAs that cannot be classified as treatment-emergent ADA because of missing titer(s) (ie, a positive ADA sample after first IMP administration in a participants with pre-existing ADA but with missing titer at this sample or at baseline).
- Participants without treatment-emergent ADA correspond to participants without treatment-emergent ADA and without any inconclusive sample after first IMP administration.
- Participants with inconclusive ADA are defined as participants which cannot irrefutably be classified as with or without treatment-emergent ADA.

ADA status at baseline, at each post-baseline visit, and overall (anytime post-baseline) will be summarized using descriptive statistics.

For ADA positive participants, titer level, and neutralizing ADA positive/negative will be summarized at each visit using descriptive statistics. In addition, the incidence of neutralizing ADA positive anytime will be summarized.

The kinetics of ADA response and the effect of positive immunogenicity findings on efficacy, safety, PK, and PD may be investigated, if applicable.

The immunogenicity analysis set will be used for immunogenicity analysis.

11. PHARMACOKINETIC ANALYSIS

Serum concentration of TEV-48574 will be listed, and summary statistics will be presented by visit. Pop-PK and PK/PD analysis will be conducted per separate analysis plans and reported separately (see Section 2.4.3).

Serum concentrations that are below the limit of quantification (BLQ) will be treated as 0 for calculation of concentration descriptive statistics.

PK collections that have an actual sampling time that deviates from predefined collection time window (+/- 3days) or that are collected post-dose will be flagged in the data listing and excluded from the calculation of serum concentration summary statistics.

Summary statistics for serum concentration of TEV-48574 will include n, arithmetic mean, SD, arithmetic CV%, median, minimum, and maximum. The summaries will be presented by indication and treatment group, by visit. Serum concentrations will be reported to 3 significant digits in summary statistics except CV%, which will be reported to 1 decimal place.

The PK analysis set will be used for the PK analyses.

12. PHARMACODYNAMIC ANALYSIS

The following pharmacodynamic biomarkers endpoints will be analyzed as part of the statistical analysis:

- Free TL1A change from baseline at each visit
- Total TL1A change from baseline at each visit

Summary statistics for free and total TL1A will be presented by visit. Actual values, change from baseline to each visit will be provided. In addition, graphs of mean change from baseline and mean percent change from baseline to each visit by treatment group will be presented.

The safety analysis set will be used for the pharmacodynamic analyses. Participants with kit mix-ups will be excluded from the analysis.

The analysis will be repeated by response at week 14 as defined by the primary endpoint.

13. PLANNED INTERIM ANALYSIS

Interim analyses for safety (periodic analysis of safety data) are planned for this trial.

The statistical analysis for the IAs will be performed by an external, independent, unblinded analysis team. Release of randomization codes to the unblinded statistician will follow SOP GBP-RD-703. At the IDMC meetings, the unblinded results will be presented and discussed in a closed session of the IDMC.

Unblinded analysis datasets, TLGs, and unblinded analysis outputs will not be disclosed to the sponsor until the trial is completed, database is locked, and randomization codes are released for final analysis. The sponsor will be informed of the IA recommendation; communication format is described in the IDMC charter.

At the safety monitoring meetings, the IDMC will review in detail any safety signal that meets the safety stopping rules described in Section 13.1, and take into account the totality of safety data. The IDMC will make a recommendation if the safety signal and the totality of the safety data warrants stopping the dose/indication/trial.

The IDMC's recommendations to modify an aspect of the trial design should maintain the Sponsor's blind.

For additional details, refer to the IDMC Charter.

13.1. Interim Safety Analyses

Periodic analysis of safety data will be done during the trial.

For each of the 4 types of AEs listed in Table 11, a Bayesian analysis using the Beta-binomial model with a non-informative *Beta (1, 1)* prior will be conducted. The posterior probability of each type of AE will be compared for each dose and the 2 dose levels combined versus placebo. If the posterior probability that the AE rate in the active group is greater than the AE rate in the placebo group reaches a posterior probability cutoff defined below, the stopping rule will be triggered. The analyses and evaluation will be conducted for UC and CD separately and across the entire trial population.

Table 11: Safety Stopping Criteria

AE Type	Posterior Probability Cutoff
Serious adverse events (aggregated)	90%
Grade 3/4 adverse events (aggregated)	90%
Opportunistic and severe and/or serious infections	85%
Anaphylaxis	85%

Details on the Beta-binomial model are provided in Appendix 1.

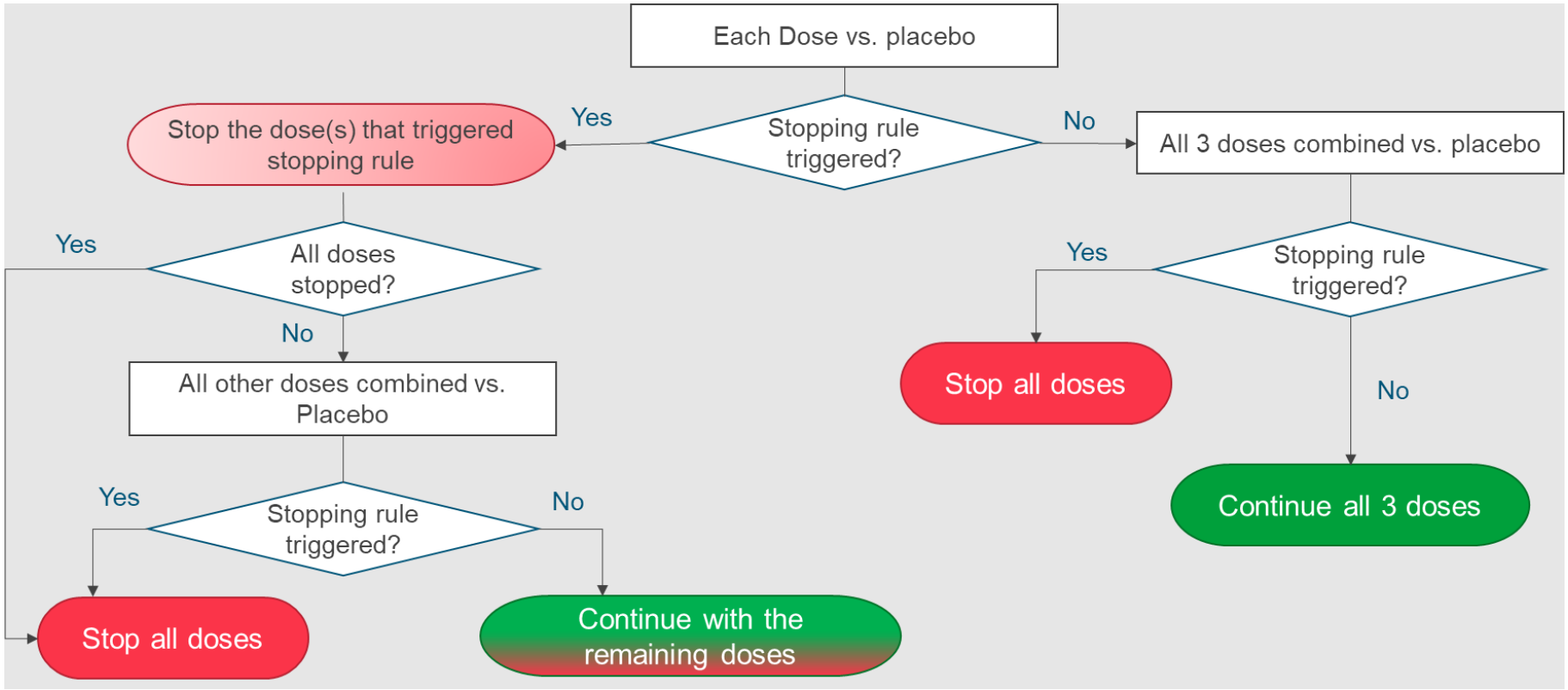
The "basic process" for applying the safety stopping rules is presented in Figure 1.

The safety interim analysis comprises of the following steps:

- Step 1: CD and UC separately, for each type of AE
  - Apply the “basic process” for each indication separately
- Step 2: Combine indications, for each type of AE (not applicable if one indication is stopped)
  - Apply the “basic process” to the combined data from both indications, excluding the dose(s) that were stopped in each indication (if applicable)
- Step 3: Detailed review
  - Review in detail any safety signal that meets the safety stopping rule, taking into account the totality of safety data
- Step 4: IDMC recommendation
  - Make a recommendation if the safety signal warrants stopping the dose/indication/trial

Steps 1 and 2 will be performed periodically by the unblinded analysis group. The Bayesian stopping rules analysis will be performed by the unblinded statistician on a monthly basis, starting after 24 participants are randomized to the trial. Steps 3 and 4 will be performed by the IDMC in case any of the stopping rules is triggered.

Figure 1: Safety Stopping Rules – Basic Process



14. STATISTICAL SOFTWARE

All data listings, summaries, and statistical analyses will be generated using SAS® version 9.4 or later unless noted otherwise.

Graphical presentation of posterior distributions and exploratory predictive modeling for response may be performed using R Version 3.5.2 or above, in the validated Advanced Computational Environment (ACE) (Redhat Enterprise Linux Server version 6.10 or higher).

15. CHANGES TO ANALYSES SPECIFIED IN THE TRIAL PROTOCOL

There are no changes to analyses specified in the trial protocol.

APPENDIX 1. BETA-BINOMIAL MODEL

Bayesian analysis using the Beta-binomial model with a non-informative Beta (1,1) prior will be performed in the following situations:

- Primary analysis (see Section 6.2.2)
- Safety stopping rules at interim safety reviews (see Section 13.1)

Efficacy Analysis – Beta Binomial Model

Using the following notation:

- p = response rate;
- d = active dose group;
- 0 = placebo group;
- i = indication
- Response = clinical remission based on MMS at week 14 for UC; endoscopic response based on SES-CD at week 14 for CD

For each indication (UC or CD) and dose (TEV-48574 450mg Q2W or TEV-48574 900mg Q2W), a Bayesian Beta-binomial model will be fit as follows:

Let n_d^i be the number of participants in the indication in the active dose group

n_0^i be the number of participants in the indication in the placebo group

Let x_d^i be the number of responders in the indication in the active dose group

x_0^i be the number of responders in the indication in the placebo group

Using a non-informative *Beta (1,1)* prior, the posterior distribution of the response rates in the active dose and placebo groups are given by

$$p_d^i \sim Beta\left(1 + x_d^i, 1 + \left(n_d^i - x_d^i\right)\right)$$

$$p_0^i \sim Beta\left(1 + x_0^i, 1 + \left(n_0^i - x_0^i\right)\right)$$

The posterior probability that the response rate in the active group is higher than the response rate in the placebo group is given by $\text{Prob}\left(p_d^i > p_0^i\right)$.

The posterior distribution of the difference between the response rate in the active dose group and the response rate in the placebo group will be estimated by obtaining the empirical difference between 1,000,000 pairs of independent samples drawn randomly from the distributions (p_d^i, p_0^i).

Let $\left\{\delta_s^{i,d}\right\}_{s=1}^{1,000,000}$ be the sampled differences. The probability $\text{Prob}\left(p_d^i > p_0^i\right)$ will be estimated by

$$\frac{1}{1{,}000{,}000}\sum_{s=1}^{1,000,000} I\left(\delta_s^{i,d} > 0\right)$$

Safety Analysis – Beta-Binomial Model

Using the following notation:

- π = AE rate;
- d = active dose group;
- 0 = placebo group;
- *i* = indication;
- k = AE type

For each of the 4 types of AEs (see Section 13.1, Table 11), dose (single or combined), indication (single or combined) a Bayesian Beta-binomial model will be fitted as follows:

Let n_d^i be the number of participants in the indication in the active dose group

n_0^i be the number of participants in the indication in the placebo group

Let $x_d^{i,k}$ be the number of participants in the indication and active dose group experiencing an AE of the selected type

$x_0^{i,k}$ be the number of participants in the indication in the placebo group experiencing an AE of the selected type

Using a non-informative *Beta (1,1)* prior, the posterior probabilities in the active dose group and placebo group are given by

$$\pi_d^{i,k} \sim Beta\left(1 + x_d^{i,k}, 1 + \left(n_d^i - x_d^{i,k}\right)\right)$$
$$\pi_0^{i,k} \sim Beta\left(1 + x_0^{i,k}, 1 + \left(n_0^i - x_0^{i,k}\right)\right)$$

The safety stopping rule is given by

$$\text{Prob}\left(\pi_d^{i,k} > \pi_0^{i,k}\right) > C^k$$

where C^k is the posterior probability cutoff for the AE type (see Section 13.1, Table 11).

The posterior probability will be estimated as described above for the efficacy analysis.

Numerical examples of the pre-defined safety stopping rules are presented in Appendix 2 .

APPENDIX 2. SAFETY STOPPING RULES

Table 12: Safety Stopping Rules – Example 1: Single TEV-48574 dose vs placebo, 18 participants, stopping criterion: 90%

		Placebo (Number of Participants With Event / Number of Participants)		
		0/9	**1/9**	**2/9**
TEV-48574 (Number of Participants With Event / Number of Participants)	**0/9**	0.500	0.237	0.106
	1/9	0.763	0.500	0.290
	2/9	0.895	0.709	0.500
	3/9	0.956	0.848	0.686
	4/9	0.984	0.929	0.824
	5/9	0.995	0.971	0.915
	6/9	0.998	0.990	0.965
	7/9	1.000	0.997	0.988
	8/9	1.000	0.999	0.997
	9/9	1.000	1.000	1.000

For each combination, the posterior probability that the event rate in the active group is higher than the event rate in the placebo group is presented

Table 13: Safety Stopping Rules – Example 2: Pooled TEV-48574 vs placebo, 36 participants, stopping criterion: 90%

		Placebo (Number of Participants With Event / Number of Participants)		
		0/9	**1/9**	**2/9**
TEV-48574 (Number of Participants With Event / Number of Participants)	**0/27**	0.263	0.064	0.014
	1/27	0.463	0.164	0.048
	2/27	0.611	0.279	0.103
	3/27	0.723	0.396	0.173
	4/27	0.804	0.507	0.257
	5/27	0.863	0.608	0.348
	6/27	0.906	0.694	0.440
	7/27	0.936	0.767	0.531
	8/27	0.958	0.827	0.615
	9/27	0.972	0.873	0.691
	10/27	0.982	0.909	0.759
	11/27	0.989	0.937	0.815
	12/27	0.993	0.957	0.862
	13/27	0.996	0.972	0.900
	14/27	0.998	0.982	0.930
	15/27	0.999	0.989	0.953

For each combination, the posterior probability that the event rate in the active group is higher than the event rate in the placebo group is presented

Table 14: Safety Stopping Rules – Example 3: Pooled TEV-48574 vs placebo, different numbers of participants, stopping criterion: 90%

Total N	Number of Placebo Participants With Event	Number of TEV-48574 Participants With Event	Probability of TEV-48574 >Placebo
36	0/7	8/29	0.912
36	0/9	7/27	0.936
36	1/7	13/29	0.914
36	1/9	10/27	0.909
72	0/16	8/56	0.919
72	0/18	7/54	0.919
72	1/16	13/56	0.916
72	1/18	11/54	0.906

APPENDIX 3. COMPARISON OF PRIMARY ANALYSIS MODEL BETWEEN PROTOCOL AMENDMENT 02 AND PROTOCOL AMENDMENT 04

This section presents a comparison of the major changes in the primary analysis model between Protocol Amendment 02 and Protocol Amendment 04.

The statistical analysis method for the primary analysis was changed from an Emax dose-response- model, which was planned when the study design incorporated 3 treatment arms of TEV-48574 active doses and a placebo arm, to a pairwise comparison approach, primarily due to the removal of 1 active treatment arm. With only 2 active dose arms and a placebo arm, the Emax model would not yield robust or precise estimates for characterizing the dose-response relationship and is no longer appropriate to assess the primary objective of the Phase 2 dose range finding study, which is to characterize the efficacy of 2 doses of TEV-48574 in patients with moderate- to-severe UC or CD.

Additionally, the previous analysis method included a Bayesian borrowing approach that relied on an assumption that both indications, UC and CD, will be concordant (drug is either effective in both indications or not effective in both indications). This is now changed to a simplified Bayesian Beta-Binomial model with no borrowing, and the model is fit separately within each indication, where each dose is compared with placebo. This approach makes minimum assumptions on the dose-response curve shape and the data structure.

Operating characteristics of this approach, for the planned sample size, were considered adequate for a Phase 2 study (See Section 2.5). For UC, for each dose the Type I error rate is 8%, and the probability to declare a dose as successful at the end of the study under the assumption of efficacy is 91%; for CD, for each dose the Type I error rate is approximately 10%, and the probability to declare a dose as successful at the end of the study under the assumption of efficacy is 87%.

Teva had considered other options for the primary analysis model, including a Bayesian hierarchical logistic regression model with dose as a categorical variable. However, similar to the original Bayesian hierarchical model, this model also makes the assumption that the drug was either effective in both indications or not effective in both indications, and it would not yield much gain in efficiency (in terms of the operating characteristics, as noted in simulations) compared with the Beta-Binomial model.

a. Comparison of design, primary analysis model, and futility IA:

Table 15: Comparison of design, primary analysis model, and futility IA between PA02 and PA04

	Old (PA02)	**New (PA04)**
Design – treatment arms	TEV-48574 1800mg Q2W, 900mg Q2W, 450mg Q2W, placebo	TEV-48574 900mg Q2W, 450mg Q2W, placebo
Model	Bayesian hierarchical Emax dose-response model	Bayesian Beta-Binomial model
Sample size	140 per indication (35 per arm per indication)	120 per indication (40 per arm per indication)
Prior	Non-informative; chosen to encompass the range of plausible values for the model parameters	Non-informative Beta (1,1) prior
Borrowing information across UC and CD	Yes, moderate borrowing	No, the model is fitted separately to each indication
Borrowing information across doses	Yes, dose-response model assumes monotonicity and structure (dose is a numerical variable in the model)	No, the model is fit separately to each dose

APPENDIX 4. DERIVATION OF CDAI SCORES

a. Definition of CDAI Score

The CDAI score is comprised of 8 items (Table 16). The total CDAI score is the sum of each item times the multiplier. If any of the subscores are missing, then the CDAI score is set to missing.

Table 16: CDAI Score

Item Description	Multiplier
Number of liquid or soft stools (each day for 7 days)	× 2
Abdominal pain, sum of 7 daily ratings (0=none, 1=mild, 2 = moderate, 3=severe)	× 5
General well-being, sum of 7 daily rating (0=generally well, 1=slightly under par, 2=poor, 3=very poor, 4=terrible)	× 7
Number of listed complications (arthritis or arthralgia, iritis or uveitis, erythema nodosum or pyoderma gangrenosum or aphthous stomatitis, anal fissure or fistula or abscess, other fistula, fever over 37.8°C [100°F])	× 20
Use of diphenoxylate or loperamide for diarrhea (0=no, 1=yes)	× 30
Abdominal mass (0=no, 2=questionable, 5=definite)	× 10
Hematocrit (males, 47-HCT [%], females, 42-HCT [%])	× 6
Body weight (1-weight/standard weight) x 100[a]	× 1

[a] if the calculated body weight item score is less than -10, set to -10
HCT = Hematocrit

b. Data Sources for the Derivation of Each CDAI Item

- Daily number of liquid or soft stools; daily abdominal pain; daily general well-being, use of diarrhea medications: Clario e-diary
- Endoscopy date: Alimentiv data
- Bowl preparation date: Clario data, if the recorded endoscopy date in Clario data matches the date in the Alimnetiv source data
  - In case of discrepancy, the day prior to endoscopy date will be assumed to be bowl preparation date
- Listed complications; abdominal mass: Clario data
- Hematocrit: Laboratory data (note: for Week 14 CDAI, hematocrit values from baseline of trial 20038 may be used, if applicable)
- Body weight: CRF data (note: for Week 14 CDAI, weight values from baseline of trial 20038 may be used, if applicable)
- Standard weight: standard height-weight table (see Table 17)
- Height: CRF

c. Window for Patient-Reported Outcomes

In general, the patient-reported outcomes of stool frequency, abdominal pain, general well-being, and use of diarrhea medications are based the e-diary data collected in a 7-day window prior to the assessment of CDAI, which may be extended to 10 days. The evaluable days for the e-diary data do not include the day of bowl preparation, the day of endoscopy, and the 2 days after endoscopy. A minimum of 3 consecutive days with completed diary entries or 4 nonconsecutive days are necessary to calculate the CDAI stool frequency, abdominal pain, and general well-being items, otherwise the score will be considered missing.

Details regarding the window for patient-reported outcomes will be provided in a separate CDAI derivation rules document that will be finalized prior to DBL.

d. Special Rules for Derivation of CDAI Score at Screening, Baseline, Weeks 2-12, and Week 14

Specific derivation rules of CDAI Score at Screening, Baseline, Weeks 2-12, and Week 14 will be provided in a separate CDAI derivation rules document that will be finalized prior to DBL.

e. Standard Height and Weight

Table 17: Standard Height and Weight Table

Actual Height cm (inches)	Standard Weight in kg Men (Pounds)	Standard Weight in kg Women (Pounds)
147.3 (58.0)		52.2 (115.0)
148.6 (58.5)		52.6 (116.0)
149.9 (59.0)		53.1 (117.0)
151.1 (59.5)		53.6 (118.3)
152.4 (60.0)		54.2 (119.5)
153.7 (60.5)		54.8 (120.8)
154.9 (61.0)		55.3 (122.0)
156.2 (61.5)		56.0 (123.5)
157.5 (62.0)	61.7 (136.0)	56.7 (125.0)
158.8 (62.5)	62.1 (137.0)	57.4 (126.5)
160.0 (63.0)	62.6 (138.0)	58.0 (128.0)
161.3 (63.5)	63.0 (139.0)	58.7 (129.5)
162.6 (64.0)	63.5 (140.0)	59.4 (131.0)
163.8 (64.5)	64.1 (141.3)	60.1 (132.5)
165.1 (65.0)	64.6 (142.5)	60.8 (134.0)
166.4 (65.5)	65.2 (143.8)	61.4 (135.5)
167.6 (66.0)	65.8 (145.0)	62.1 (137.0)
168.9 (66.5)	66.4 (146.5)	62.8 (138.5)

Table 17: Standard Height and Weight Table (Continued)

Actual Height cm (inches)	Standard Weight in kg Men (Pounds)	Standard Weight in kg Women (Pounds)
170.2 (67.0)	67.1 (148.0)	63.5 (140.0)
171.5 (67.5)	67.8 (149.5)	64.2 (141.5)
172.7 (68.0)	68.5 (151.0)	64.9 (143.0)
174.0 (68.5)	69.2 (152.5)	65.5 (144.5)
175.3 (69.0)	69.8 (154.0)	66.2 (146.0)
176.5 (69.5)	70.5 (155.5)	66.9 (147.5)
177.8 (70.0)	71.2 (157.0)	67.6 (149.0)
179.1 (70.5)	71.9 (158.5)	68.3 (150.5)
180.3 (71.0)	72.6 (160.0)	68.9 (152.0)
181.6 (71.5)	73.4 (161.8)	69.6 (153.5)
182.9 (72.0)	74.1 (163.5)	70.3 (155.0)
184.2 (72.5)	75.0 (165.3)	
185.4 (73.0)	75.7 (167.0)	*Height in shoes with one-inch heels
186.7 (73.5)	76.6 (169.0)	
188.0 (74.0)	77.5 (171.0)	*Indoor clothing weighing 5 pounds for men and 3 pounds for women
189.2 (74.5)	78.4 (172.8)	
190.5 (75.0)	79.1 (174.5)	
191.8 (75.5)	80.2 (176.8)	*Centimeters x 0.3937 = inches
193.0 (76.0)	81.2 (179.0)	*Pounds x 0.4536 = kilograms

APPENDIX 5. PRO2-CD SCORE

Table 18: Definition of PRO2-CD Score

Score	Data	Derivation
PRO2-CD Stool Frequency Subscore	Mean number of liquid or soft stools in 7-day period	Mean value <0.5 → subscore=0; Mean value ≥0.5 and <1.5 → subscore =1; Mean value ≥1.5 and <2.5 → subscore = 2; Mean value ≥2.5 → subscore =3
PRO2-CD Abdominal Pain Subscore	Mean abdominal pain in 7-day period	
PRO2-CD Score	Sum of stool frequency and abdominal pain subscores	

Derivation of PRO2-CD Score:

- Calculate the stool frequency and abdominal pain mean values at: derived baseline, weeks 2 to 12, and week 14 using the rules as described in Appendix 4 for CDAI.
- Derive the subscores (range 0-3) as described in Table 18.
- PRO2-CD is the sum of the stool frequency and abdominal pain subscores.

Note that the binary efficacy endpoints based on PRO2-CD (see Table 5) use cutoffs for the mean values and not the subscores.

APPENDIX 6. CORTICOSTERIODS - EQUIVALENT PREDNISONE DOSE

The equivalent prednisone dose for different oral corticosteroids, is provided in Table 19.

Table 19: Equivalent Prednisone Dose

Corticosteroid	Equivalent Dose (mg)
Cortisone	25
Hydrocortisone	20
Prednisone	5
Prednisolone	5
Triamcinolone	4.0
Methylprednisolone	4.0