STATISTICAL ANALYSIS PLAN	
Study Code	D5271C00001 (Legacy # 3150-301-008)
Edition Number	3.0
Date	28-Mar-2023

FOR STAGE 1 (PHASE 2B) OF THE STUDY

A 52-Week, Multicenter, Randomized, Double-blind, Placebo and Active-Controlled, Operationally Seamless Phase 2b/3, Parallel-group Study to Assess the Efficacy and Safety of Brazikumab in Participants with Moderately to Severely Active Crohn's Disease (INTREPID Lead-In)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation/Term	Definition
ADA	Anti-drug antibody
AE	Adverse event
AESI	Adverse event of special interest
ALT	Alanine aminotransferase
AP	Abdominal pain
AST	Aspartate aminotransferase
ATC	Anatomical therapeutic chemical
BM	Biomarker
BM-	Biomarker serum IL-22 concentrations below a pre-established cutoff
BM+	Biomarker serum IL-22 concentrations at or above a pre-established cutoff
BMD	Bowel movement diary
BMI	Body mass index
BSFS	Bristol stool form scale
BUN	blood urea nitrogen
CD	Crohn's disease
CDAI	Crohn's disease activity index
CDL	Clinical data lock
CI	Confidence interval
CMH	Cochran Mantel-Haenszel
CRP	C-reactive protein
CS	Corticosteroid
CSP	Clinical study protocol
CSR	Clinical study report
CV	Coefficient of Variation

Abbreviation/Term	Definition
E/D	Early Study Intervention Discontinuation
ECG	Electrocardiogram, electrocardiographic
eCRF	Electronic case report form
eDiary	Electronic diary
eGFR	Estimated glomerular filtration rate
EIM	Extraintestinal manifestation
EQ-5D-5L	European Quality of Life- 5 Dimensions
ET	Early termination
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue Scale
FAS	Full Analysis Set (intent to treat)
FCP	Fecal calprotectin
FSH	Follicle stimulating hormone
GCP	Good clinical practice
GGT	Gamma-glutamyl transferase
hCG	Human chorionic gonadotropin
HCP	Health care professional
HIV	Human immunodeficiency virus
HRQoL	Health-related quality of life
IA	Interim analysis
IBD	Inflammatory Bowel Disease
IBDQ	Inflammatory Bowel Disease Questionnaire
IBW	Ideal Body Weight
ICF	Informed consent form
ICH	International Conference on Harmonization
IgM	Immunoglobulin M
IL12/23	Interleukin-12/23

Abbreviation/Term	Definition
IL-22	Interleukin-22
IP	Investigation product
IPD	Important protocol deviations
ITT	Intent-to-treat
IV	Intravenous
IXRS	Interactive voice/web response system
JAK	Janus kinase
LLN	Lower limit of normal
LOCF	Last observation carried forward
LS	Least squares
LSF	Loose stool frequency
MAR	Missing at random
Max	Maximum
MCH	Mean corpuscular hemoglobin
MCSc	Mental Component Score
MCV	Mean corpuscular volume
MedDRA	Medical Dictionary for Regulatory Activities
Min	Minimum
MMRM	Mixed Model for Repeated Measures
MRD	Minimum Required Dilution
n	Number of participants
NRI	Non-responder imputation
NSAID	Non-steroidal anti-inflammatory drug
NRS	Numerical Rating Scale
OLE	Open label extension
PCS	Potentially Clinically Significant

Abbreviation/Term	Definition
PCSc	Physical Component Score
PK	Pharmacokinetic
PP	Per protocol
PRO	Patient-reported outcome
PT	Preferred term
Q1	Lower quartile
Q3	Upper quartile
QRS	Time from beginning of the Q wave to end of the S wave in heart's electrical cycle
QT	Time from beginning of the Q wave to end of the T wave in heart's electrical cycle
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using the Bazett formula ($QTcB = QT/(RR)^{1/2}$)
QTcF	QT interval corrected for heart rate using the Fridericia formula ($QTcF = QT/(RR)^{1/3}$)
RBC	Red blood cell
RR	Respiratory rate
SAE	Serious adverse event
SAP	Statistical analysis plan
SC	Subcutaneous
SD	Standard deviation
SE	Standard error
SES-CD	Simple Endoscopic Score for Crohn's Disease
SF-36v2	Short Form 36-Item Health Survey Version 2
SI	Le Système International d'Unités (International System of Units)
SoA	Schedule of Actiivities
SOC	System organ class
SV2	Screening Visit 2
TB	Tuberculosis

Abbreviation/Term	Definition
TE	Treatment emergent
TNF	Tumor necrosis factor
TPV	Third party vendor
ULN	Upper limit of normal
US	United States
VAS	Visual analog score
WBC	White blood cell
WD	Weighted difference
WHO	World Health Organization

AMENDMENT HISTORY

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Entire document	9 Nov 2020 (Version 1.0)	Original SAP as transferred from Allergan with change of sponsor to AstraZeneca and change of study number to AstraZeneca study number.	N/A	N/A
Entire document	10 Nov 2021 (version 2.0)	Updated SAP to reflect Protocol Amendment 4 Version 5 changes	Yes	N/A
Entire document	28 Mar 2023 (version 3.0)	Aligned descriptions of tables, listings, and figures to align with AZSOL standards.	Yes	N/A
Entire document		Replaced “protocol” with “CSP”, “Phase 2b” with “Stage 1”, “patients”, and “subjects” with “participants”.	Yes	Made language consistent
Entire document		Replaced “analyse” with “analyze”	Yes	Made American spelling consistent
Entire document		Removed references to eCRF entries	N/A	To avoid SAP updates if CRF is updated.
Section 3.2, Section 4.1.2.1		Added ADA Evaluable Analysis Set. Described further the FAS, Safety Analysis set, and PK Analysis Set	Yes	To account for erroneously treated participants, and participants enrolled prior to CSP Amendment 4.
Section 3.3		Added rules for number of decimal places.	N/A	To clarify rounding rules of presented estimates
Section 3.3.1.2		Updated baseline definition of CDAI PRO items, and added baseline definition of FCP.	N/A	To clarify how baseline will be derived for these endpoints.
Section 3.3.1.2		Added derivation of Change from Baseline, and Percent Changed from Baseline	N/A	To clarify how these variables will be derived.
Section 3.3.1.3		Updated study periods	Yes	To account for early terminations and improved clarity.
Section 3.3.2.2		Updated analysis windows of Week 52, End of Treatment, and Follow-up Visits.	Yes	To account for early terminations and improved clarity.
Section 3.3.2.3		Updated handling of early discontinuation visits.	N/A	To clarify how these assessments will be handled.
Section 3.3.2.4		Added text to handle late ileocolonoscopy assessments at Week 12	N/A	To clarify how these assessments will be handled.

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Section 3.3.4		Updated text.	Yes	To align with definition of FAS and improve clarity.
Section 3.3.5		Corrected placement of abbreviation after first usage	N/A	Improved readability.
Sections 4.1		Deleted all references to presentations “by BM status”.	Yes	Aligned with CSP
Section 4.1.1.1		Updated definition of study intervention completion. Added definition of study completion.	N/A	To clarify how these variables will be derived.
Section 4.1.1.2		Updated categories of disposition	Yes	To clarify how these categories will be presented.
Section 4.1.1.2		Replaced the summary of COVID-19 study disruption with summaries of global/country situation	Yes	To clarify how these categories will be presented.
Section 4.1.3.2		Added IPD categories	N/A	Aligned categories with Protocol Deviation Plan.
Section 4.1.5.1		Removed BMI calculation.		Derived in database.
Section 4.1.5.1		Moved Change from Baseline definition to Section 3.3.1.2	N/A	Improved readability of the document.
Section 4.1.6.1		Replaced perianal fistula with fistula. Removed substance use. Added prior use of JAK inhibitors. Removed evening diary, BMD, and site visit instruments. Added list of Immunomodulators, JAK inhibitors and ATC codes.	Yes	To clarify which categories will be presented at disease characteristics and how those will be identified.
Section 4.1.7.1		Removed “substance use” from Crohn’s disease medical history	Yes	Smoking and alcohol status were already shown in the tables of baseline disease characteristics
Section 4.1.7.2		Removed summary of concurrent procedures due to CD.	Yes	Updated to simplify reporting and align with AZSOL template..
Section 4.1.8.1		Updated definition of concomitant medication	N/A	To clarify that concomitant medication starting 4 weeks after the last dose of IP will not be included in the summary, but will be listed.
Section 4.1.8.2		Replaced “preferred term” with “generic drug name”	Yes	Updated to align with AZSOL template
Section 4.1.8.2		Updated list of tables that will be provided.	N/A	To clarify which tables will be created.

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Section 4.1.9.2		Added categorical summary of number and percentage of doses at Week 0, 4, 8, and during the induction period. Added listing of IP administration.	N/A	To clarify which tables will be created.
Section 4.2		Added the paragraph "Parametric tests are described for each endpoint analysis. The data will undergo appropriate tests of the normality assumptions and if necessary, data will be transformed or non-parametric substitutes will be used for analysis."	Yes	To add flexibility if statistical assumptions of planned analyses are not met.
Section 4.2 Table 4-5		Deleted ": SES-CD total score of 0-2, OR SES-CD total score of ≤ 4 and at least 2 point reduction from Baseline with no subscore > 1" from 1st line of objective 3"	Yes	Phrase was added in error.
Section 4.2, Table 4-5		Removed Objective 10 and all associated endpoints	No	All endpoints are still included when evaluated for participants in the FAS who take CS at baseline.
Section 4.2, Table 4-5		Updated primary symptom remission (associated with Objective 12 (previous 13)) to only reference Week 12	Yes	Aligned with CSP.
Section 4.2, Table 4-5		Updated clinical response (associated with Objective 13 (previous 14)) to only reference Week 12	Yes	Aligned with CSP.
Section 4.2, Table 4-5		Removed all endpoints from Objective 20 except: • CDAI response • CDAI remission • Clinical response • Clinical remission • CDAI remission and endoscopic response • Clinical remission and endscopic response	Yes	Aligned with CSP.
Section 4.2.1.2, Appendix 7.1		Updated text so that the whole derivation of CDAI score lies within Section 4.2.1.2. Corrected formula of IBW so that it is based on height in cm.	N/A	Improved readability. Height is standardized to cm in database.
Section 4.2.1.3		Added the sentence "NRI for intercurrent events will be applied prior to missing data handling."	N/A	Added to clarify order of programming.

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Section 4.2.1.4		Corrected references from "model" to "test" Corrected typos in formulae	N/A	CMH is NPR test, rather than a model. To ensure correct implementation of the test
Section 4.2.1.4.1		Updated description of what outputs should be presented for the primary analysis. Updated description of NRI table.	Yes	Clarified which categories, and estimates will be presented.
Section 4.2.1.4.1		Removed "Pair-wise Mantel Haenzel odds ratios"	Yes	The comparison is rate difference.
Section 4.2.1.4.1		Remove "by BM status" in the primary analysis presentation	Yes	Removed duplicate analysis of BM status.
Section 4.2.1.5		Updated text. Added reference.	N/A	Clarified how the logistic regression model will be set up.
Section 4.2.1.6, Section 4.2.2.6, Section 4.2.3.5		removed Mixed strategy	N/A	No longer needed.
Section 4.2.1.7		Add the sentence "A "BY" statement will be added to the model specification in each case to create the subgroups."	Yes	Clarifying language added for specification of the subgroups.
Section 4.2.1.7		Add the sentence "For testing of "prior biological use" and "Current CS use at randomization" as subgroups, the subgroup variable will not be included as a covariate."	Yes	Clarified that randomization strata will not be included in the model as both a subgroup and a covariate at the same time.
Section 4.2.2.2		Removed the sentence referring to LOCF.	N/A	It is not deemed appropriate to carry forward missing SES-CD scores.
Section 4.2.2.3		Updated text referring to missing data handling	N/A	Clarified that it is the vendor handling the SES-CD data and explained the rule.
Section 4.2.3.1		Added composite strategy	N/A	To clarify how those endpoints will be handled.
Section 4.2.5.2		Added "The cutoff used for BM status classification in the rest of this SAP is the value that maximizes the differential effect of CDAI remission at Week 12."	Yes	Added to clarify the derivation of the cutoff point used for BM Status throughout the rest of the SAP
Section 4.2.5.2		Removed text referring to graphs of combined CDAI remission and endoscopic response	N/A	Phrase was added in error.

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Section 4.2.5.4		Updated text to conduct the analysis on CDAI remission, and endoscopic response at Week 12, for each of brazikumab doses, and for the pooled brazikumab doses versus placebo.	N/A	To clarify which analyses will be conducted.
Section 4.2.5.4.1		Removed text referring to graphs of combined CDAI remission and endoscopic response	N/A	Phrase was added in error.
Section 4.2.6		Delete duplicate line “Clinical response for Weeks 2,4,8,12,40, and 50”	Yes	Deleted duplicate endpoint
Section 4.2.6		Removed references to endpoints associated with Objective 10	No	All endpoints are still included when evaluated for participants in the FAS who take CS at baseline.
Section 4.2.6		Delete composite endpoints; “CDAI remission and endoscopic response at Week 52”; “CDAI remission and endoscopic remission” endpoints; “clinical remission and endscopic response” endpoints; “both clinical remission and endoscopic remission”; “primary symptom response” for all but Week 12 endpoint; “clinical response” for all but Week 12.	Yes	Aligned with CSP.
Section 4.2.6.1		In the definitions of composite clinical response and composite clinical remission, replace “CDAI” with “clinical”	Yes	Fixed an error.
Section 4.2.6.1		Added composite strategy	N/A	To clarify how those endpoints will be handled.
Section 4.2.6.2		Updated derivation of CS-free endpoints.	Yes	To clarify how those endpoints will be derived.
Section 4.2.7.3		Updated text referring to missing data handling.	N/A	To clarify how missing data will be handled.
Section 4.2.7.4		Removed reference to random effects, and associated variance components.	N/A	There are no random effects in the model specification.
Section 4.2.7.4		Updated specification of model and variance covariance structure.	N/A	To ensure correct implementation of the model.
Section 4.2.7.4		Replace “heterogeeous” with “heterogeneous”	Yes	Spelling error.

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Section 4.2.7.4.1		Removed the COVTEST option.	N/A	This was added in error.
Section 4.2.7.5		Updated specification of model.	N/A	To ensure correct implementation of the model.
Section 4.2.8		Merged previous Section 4.2.10 with Section 4.2.8. Updated text throughout the section.	N/A	Improved readability of the document. Clarified how the variables will be derived, specified handling of missing data, and conduct of analyses.
Section 4.2.9		Updated text throughout the section.	N/A	Improved readability of the document.
Section 4.2.9.3		Removed missing data handling of SF36 questions.	N/A	Missing questions are handled by vendor.
Section 4.2.9.4		Specified that the endpoints will be analyzed via a MMRM.	Yes	To ensure correct implementation of the analysis.
Section 4.3		Updated text throughout the section.	N/A	Improved readability of the document.
Section 4.4		Removed text referring to PK variables	N/A	Detailed analysis of PK may be conducted in separate PK analysis plan and reported outside of CSR.
Section 4.5		Updated text throughout the section.	N/A	Aligned with AZ 'Standards for reporting ADA data in a Clinical Study Report or High Level Document'
Section 4.6		Remove presentation by BM+/BM- throughout	N/A	Removed to simplify outputs
Section 4.6		Updated text throughout the section.	N/A	Improved readability of the document.
Section 4.6.1.2		Removed exposure by amount, number of treatments, number and percentage receiving entire treatment.	N/A	Simplify and align with AZSOL
Section 4.6.2.1		Removed definition of Treatment Emergent AE	N/A	Aligned with AZSOL templates
Section 4.6.2.1		Specified reporting periods of AEs.	N/A	To enable reporting of AEs as appropriate
Section 4.6.2.2		Added new tables and clarified language of table contents	N/A	Aligned with AZSOL templates
Section 4.6.2.2		Updated list of summaries to be created	N/A	To clarify which tables will be created.
Section 4.6.2.2		Replaced "event rate" with "adverse events per 100 subject year"	N/A	Aligned with AZSOL template.

CATEGORY Change refers to:	Date	Description of change	In line with CSP?	Rationale
Section 4.6.3.1		Removed text referring to AEs due to Laboratory values and vitals signs.	N/A	These will be handled and reported as described in Section 4.6.2.
Section 4.6.6		Removed reference summary of results	N/A	Physical exam results will only be listed
Section 4.6.7.2		Deleted table of AE for subjects with PCS vital signs	N/A	Not required by AZSOL
Section 4.6.8.1		Updated table 4-12, thresholds of high value, high increase of QT, QTcF, QTcB. Also, added ECG heart rate in the table.	N/A	Adopted a more conservative rule.
Section 4.6.8.2		Deleted table of AE for subjects with PCS ECG results and shift table.	N/A	Not required by AZSOL
Section 4.6.9.1		Updated potential Hy's law definition to "AST>=3xULN and TBL ≥ 2x ULN, where at least one time Total Bilirubin ≥ 2x ULN occurred after the first occurrence of AST ≥ 3x ULN or ALT ≥ 3x ULN after start of treatment with investigational product."	Yes	Aligned potential Hy's law definition to AZSOL template.
Section 5		Updated wording referring to Interim (primary) analysis.	Yes	Clarified that at the time of Interim analysis, all data will be analyzed.
Appendix 7.5.1		For item CDAI04 (Fever), changed answers so that Yes = [1] and No = [0].	Yes	Corrected error in original table from CSP.

1 INTRODUCTION

Clinical Study Protocol (CSP) D5271C00001 (Legacy #3150-301-008) is planned as a global, multicenter (up to 400 sites), randomized, double-blind, placebo- or active-controlled, parallel-group, operationally seamless, Phase 2b/3 52-week study.

This statistical analysis plan (SAP) describes the statistical analysis methods for the first stage (Stage 1, also known as Phase 2b study stage) of an operationally seamless Phase 2b/3 clinical study protocol D5271C00001 (Legacy # 3150-301-008). Throughout this SAP, the term Stage 1 will be used to refer to the Phase 2b dose-ranging, efficacy and safety clinical trial phase in CSP D5271C00001 (Legacy # 3150-301-008).

This Stage 1 study SAP details technical specifications of the Stage 1 statistical analyses of the efficacy and safety data as per D5271C0001 clinical study protocol (Amendment 5 version 6, dated 07 December 2021). Specification of tables, figures, and data listings are provided in a separate document. The SAP for pharmacokinetic/pharmacodynamic data will be prepared separately.

The primary analyses In Stage 1 will be performed when all participants have completed their Week 12 visit or have discontinued early. Thus, Week 12 is the primary timepoint in the study as opposed to the final timepoint at Week 52. Accordingly, the pre-specified primary analyses fulfil the ICH E9 definition of an interim analysis. The study will be unblinded at this time, and the results from the primary analysis will serve as final results. Therefore, there will be 2 clinical database locks (CDL) in the study, at the time of the primary analyses when all participants complete Week 12 (or discontinue early in the Induction Period) and after the last participant completes last visit in the study (this could be Early Termination, Week 52, or after the safety follow-up period).

The interim (primary) analyses for this Stage 1 study based on Week 12 data and final analysis based on Week 52 data are described in this SAP. All available data from participants who have completed Week 52 or would have completed Week 52 if they hadn't discontinued, will be analyzed in the interim (primary) analyses. A Study Integrity Plan has been developed to describe the strategy for protecting the integrity of the study until the final database lock.

Approximately 240 participants will be randomized in the Stage 1 portion of the study. Prior to Amendment 4, 17 participants were enrolled in one of 3 treatment groups (2 brazikumab doses and humira) or placebo, as described below.

After Amendment 4, participants will be randomized in a 5:5:3 ratio (85 participants for each of the two brazikumab dose groups, and 51 participants for the placebo group). Participants will be stratified according to status of prior biologic use, defined as biologic naïve/non-refractory or biologic refractory/intolerant, and current corticosteroid (CS) use at randomization (Yes or No).

Following a Screening Period of up to 5 weeks, the study duration is up to 66 weeks, consisting of a 52-week Double-Blind Treatment Period, and an 18-week post-treatment Safety Follow-up Period (starting at Week 48). For participants who enroll into an open-label extension study of brazikumab (Study D5271C00002 (Legacy # 3150-303-008)) after the 52-week Double-Blind Treatment Period in Stage 1, the Safety Follow-up Period will not be applicable.

There are 3 treatment groups in Stage 1:

- Brazikumab high dose:
  IV brazikumab 1440 mg on Days 1, 29, and 57, followed by SC brazikumab 240 mg on Day 85 and every 4 weeks through Week 48
- Brazikumab low dose:
  IV brazikumab 720 mg on Days 1, 29, and 57, followed by SC brazikumab 240 mg on Day 85 and every 4 weeks through Week 48
- Placebo:
  IV placebo on Days 1, 29, and 57, followed by SC placebo on Day 85 and every 4 weeks through Week 48

Prior to implementation of Amendment 4, there was an additional treatment group:

- Humira®:

  SC injection 160 mg on Day 1, 80 mg on Day 15, followed by 40 mg on Day 29 and every 2 weeks through Week 50

The schedule of assessments for Stage 1 can be found in the CSP Section 1.3, Tables 1, 2, 3, and 4.

2 CHANGES TO PROTOCOL PLANNED ANALYSES

CHANGE	JUSTIFICATION
The following CSP specified CS-free endpoints are removed when evaluated for the FAS: CDAI remission, CDAI response, endoscopic response, endoscopic remission, SES-CD total score of 0-2, and clinical remission at Week 52. All endpoints are still included when evaluated for participants in the FAS who take CS at baseline.	It is not considered relevant to evaluate the CS-free endpoints for patients who do not take CS at baseline.

3 DATA ANALYSIS CONSIDERATIONS

3.1 Timing of Analyses

An interim (primary) analysis will be performed after all randomized participants have completed Week 12 or discontinued early in the Induction Period. The final analysis will be undertaken after the last participant completes the last visit in the study (this could be Early Termination, Week 52, or after the safety follow-up period). All analyses described in this SAP will be included in the clinical study report (CSR).

3.2 Analysis Populations

The following populations are defined for Stage 1 (Table 3-1):

Table 3-1 Populations for Analysis

Population/Analysis set	Description
Screened analysis set	This includes all participants who signed informed consent form for the study.
Full Analysis Set	This population includes all participants who are randomized to a Stage 1 treatment group (as per the ITT principle) after implementation of CSP Amendment 4. Participants will be summarized according to the randomized study intervention.

Safety Analysis Set	This includes all participants who receive ≥ 1 administration of Stage 1 study intervention[1], including all participants enrolled prior to Amendment 4. Erroneously treated participants (e.g., those randomized to treatment A but actually given treatment B) are accounted for in the treatment group of the treatment they actually received. A participant who has on one or several occasions received active study intervention is classified as active and is accounted for in the active study intervention treatment group. A participant who has received both 720 mg and 1440 mg active dose regimens is classified as the higher active dose regimen.[2]
PK Analysis Set	The PK population includes all participants in the Safety Analysis Set who have at least 1 PK sample containing detectable brazikumab concentrations. Participants will be summarized according to the actual study treatment group.
ADA Evaluable Analysis Set	This includes Safety analysis set participants having non-missing baseline ADA result and at least one non-missing post baseline ADA result. Participants will be summarized according to the actual study treatment group.

ADA = Anti-drug antibodies; ICF = Informed Consent Form; ITT = intent to treat; PK = pharmacokinetics

[1] Participants who only received Humira® will only be included in the listings.

[2] "Higher active dose" refers to the higher IV dose (1440 mg).

It should be noted that the Full Analysis Set (FAS) includes all participants enrolled after the CSP amendment 4, excluding participants randomized prior to the CSP amendment 4, which are only in the INFORM legacy database. This means that INFORM data will not be used for efficacy analyses, but only RAVE data.

3.3 General Considerations

Statistical analysis will be performed using the SAS Software, Version 9.4 or higher.

Efficacy analyses will be performed only for the participants enrolled after the date that Amendment 4 of D5271C00001 (Legacy #3150-301-008) was implemented. Efficacy data for participants who were enrolled prior to CSP Amendment 4 will be listed only; all efficacy analyses will be performed using the FAS. All Pre-Amendment 4 safety data will be summarized together with the safety data collected after CSP Amendment 4. Safety analyses will be performed using the Safety analysis set; however, safety results for participants receiving Humira® will be listed only.

Study-related raw data for enrolled participants, including derived data, will be presented in data listings. Unless otherwise stated, listings will be sorted for presentation in order of

treatment group, study site, participant number, date of event, and parameter name (if applicable).

Unless otherwise specified, descriptive statistics for continuous variables include the number of participants (n), mean, standard deviation (SD), median, 1st (Q1) and 3rd (Q3) quartiles, minimum (Min), and maximum (Max). Summary statistics for categorical variables include the number of participants (n) and percent (%).

Unless otherwise specified, presentation will be up to 3 decimal places and as follows: Minimum, and maximum will be presented to the same number of decimal places as in the recorded data. Mean, median, Q1, Q3, and CI will be displayed with 1 more decimal place than the recorded data. SD and standard error (SE) will be presented to 2 more decimal places than the recorded data. Percentages will be rounded to 1 decimal place, except 0 and 100%, which will be displayed without any decimal places. P-values will be reported to 3 decimal places with values less than 0.001 displayed as <0.001.

All efficacy and safety data will be analyzed based on the analysis visits. The details of analysis visit windowing rules will be described in Section 3.3.2.2. Tabular summaries will be presented by treatment group.

In general, unless otherwise stated, efficacy analyses will be conducted with a two-sided test at a significance level of $\alpha = 0.05$. All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise. There will be no multiplicity control across endpoints; nominal p-values will be provided as descriptive measures to aid in interpretation of the analysis results and should not be interpreted as statistical significance.

3.3.1 General Study Level Definitions

3.3.1.1 Study Day 1

Day 1 will be defined as the date of first dose of study intervention. If a participant is randomized but discontinues participation without receiving any study intervention, then their Study Day 1 will be undefined.

3.3.1.2 Baseline

Unless otherwise specified, baseline is defined as the last non-missing assessment before the first dose of study intervention. Baseline for IL-22, CDAI and individual CDAI components, and Fecal calprotectin (FCP) variables are defined below:

- For IL-22, baseline is defined as the value at Screening Visit 2 (SV2).
- For CDAI and all CDAI individual components, Baseline is defined as the SV2 assessment performed prior to initiation of bowel prep for ileocolonoscopy.
- The Baseline patient-reported outcome (PRO) items of the CDAI (LSF, AP, General Well-being, and Temperature) will be collected during the 7 days prior to the physician CDAI assessments (presence of clinical signs, antidiarrheal use, and presence of abdominal mass) and prior to initiation of bowel preparation for the Screening ileocolonoscopy.

- For hematocrit, baseline is defined as the value at Screening Visit 2 (SV2), if it is not available (e.g., hemolyzed sample), Screening Visit 1 will be used to calculate CDAI eligibility and Baseline CDAI scores, respectively. In this case eligibility and Baseline CDAI may be different.
- For FCP, baseline is defined as the value at SV2 and prior to initiation of bowel preparation for the Screening ileocolonoscopy.

Change from Baseline is defined as (Value at Visit X – Value at Baseline).

Percent Change from Baseline is defined as (Value at Visit X – Value at Baseline)/(Value at Baseline)×100.

Missing data handling for the individual CDAI components is explained in Section 4.2.1.3.

3.3.1.3 Study Periods

The study periods defined in this section are different from the safety reporting periods for Adverse Events (AE) defined in Section 4.6.2.1.

Screening Period: -35 to -1 days (Screening Visits 1, 2 and Ileocolonoscopy)

Induction Period: Day 1 to Day 98, or day of early termination prior to Day 98

Maintenance Period: Day 99 to analysis window Week 52, or day of early termination or rolling over to OLE on or later than Day 99

Double-Blind Treatment Period: Induction through Maintenance Periods

Safety Follow-up Period: 18 weeks after the last dose of brazikumab or placebo (Follow-up Visit 1 [8 weeks after last dose] and Follow-up Visit 2 [18 weeks after last dose])

3.3.2 Visit Window

3.3.2.1 Study day

For assignment of data to analysis windows, study days will be defined in reference to Study Day 1 as defined in Section 3.3.1.1.

For any analysis visit time windows, if the assessment date is on or after the date of first dose of study intervention, the study day is calculated by:

assessment date – date of first dose of study intervention + 1.

If the assessment date is before the date of first dose of study intervention, the study day is calculated by:

assessment date – date of first dose of study intervention.

Therefore, a negative day indicates a day before the start dose of the study intervention.

If the assessment date is unavailable, the visit date will be used instead.

3.3.2.2 Conventions for the visit window

Table 3-2 presents the analysis windows assigned for efficacy (except daily data parameters) and safety analyses and the corresponding range of treatment days (window) during which data for a particular nominal visit may be collected. The analysis windows will be calculated from the date of first dose of study intervention. These are analysis windows that contain, but are not the same as, the actual visit windows. Each analysis window is equidistant between the neighboring Scheduled Visit Days. For example Week 2 = Day 15 and Week 4 = Day 29 (Table 3-2). The center of the analysis window is calculated as the average between the two Visit Days: (15+29)/2= 22. Therefore the end of the Week 2 analysis window is Day 21 and the beginning of the analysis window for Week 4 is Day 22.

Similarly, for Follow-up Visit 1 and Follow-up Visit 2 scheduled at (last dose + 8 weeks) and (last dose + 18 weeks), the middlepoint is (last dose + 13 weeks), which is considered as the limit of the analysis windows during follow-up.

Table 3-2 Analysis Windows

Derived Visit	***Scheduled Visit Day***	***Analysis Window***
Screening Visit 1	-35	[-35,-15]
Screening Visit 2	-14	[-14,-1]
Week 0	Day 1[a]	Days =1
Week 2	Day 15	Days [2,21]
Week 4	Day 29	Days [22,42]
Week 8	Day 57	Days [43,70]
Week 12	Day 85	Days [71,98]
Week 16	Day 113	Days [99,126]
Week 20	Day 141	Days [127, 154]
Week 24	Days 169	Days [155, 182]
Week 28	Day 197	Days [183,210]
Week 32	Day 225	Days [211,238]
Week 36	Day 253	Days [239, 266]
Week 40	Day 281	Days [267,294]
Week 44	Day 309	Days [295,322]
Week 48	Day 337	Days [323,350]
Week 52	Day 365	Day ≥ 351 to (last dose + 6 weeks)
End of Treatment Period[b]	The earliest of Week 52 and Early Termination Visit	
Follow-up Visit 1[c]	(last dose + 6 weeks + 1) to (last dose + 13 weeks)	

Follow-up Visit 2[c]	(last dose + 13 weeks + 1) to End of Study

[a] Relative to the date of the first dose of study intervention in the Induction Period. Day 1 = the date of the first dose of study intervention in the Induction Period. There is no Day 0.
[b] Presented in analysis tables for safety parameters, including but not limited to electrocardiograms, clinical laboratory values, and vital signs.
[c] Follow-up Visits will only be carried out for participants who do not join the open label extension study.

3.3.2.3 Early discontinuation visits

Early discontinuation visits for either discontinuation of study intervention or early discontinuation from study will be mapped to analysis windows following the conventions in Table 3-2. The closest assessment to the Scheduled Visit day should be used regardless if it is scheduled or early discontinuation. Both early discontinuation and scheduled visits will be used in case of multiple observations within an analysis visit window and same rules as in Section 3.3.2.4 will be applied.

3.3.2.4 Multiple observations

All observations will be assigned to their corresponding analysis window following the conventions in Table 3-2. If multiple assessments are recorded in a single analysis window, the following rules will be followed unless otherwise specified:

- If there are 2 or more observations within the screening analysis window, the last non-missing assessment before the the first study intervention will be used.
- If there are 2 or more observations within the same analysis window, then the observation closest to the Visit where study intervention was given will be used in the analysis.
- If there are two Visits with study intervention, the observation closest to the planned protocol Scheduled Visit day will be used. The other observation(s) will not be used for analysis but will be included in listings.
- If 2 observations are equidistant from the study day when the study intervention was given, the observation with the earlier collection date will be used in the analysis.
- At Week 2 and Week 52, where study intervention is not planned, then the observation closest to the Scheduled Visit day will be used in the analysis.
- At Week 12, ileocolonoscopy assessments related to the Induction Period will not be used in the analysis if they take place more than 7 days after receipt of SC study intervention.

If the physician CDAI assessments (presence of clinical signs, antidiarrheal use, and presence of abdominal mass) and ileocolonoscopy are done at same visit, use the PRO assessments that are prior to the bowel prep. If there are multiple CDAI total score assessments during the Screening Period, the assessment that is closest to Screening Visit 2 and prior to bowel prep before the Screening ileocolonoscopy will be used. Once the assessment date is determined, the diary (PRO) data for the seven days prior to initiation of the bowel prep will be used to calculate the CDAI score.

3.3.3 Handling of Unscheduled Visits

If the study intervention cannot be administered during a scheduled study visit, the rescheduled dose will be administered as an unscheduled visit.

Post-baseline unscheduled visits (including early termination visits) will be mapped to a visit based on the analysis windows in Table 3-2. Both unscheduled and scheduled visits will be used in case of multiple observations within an analysis visit window and same rules as in Section 3.3.2.4 will be applied.

Data from unscheduled visits will be presented in data listings.

3.3.4 Multiplicity/Multiple Comparisons

The null hypothesis is that CDAI remission rate (primary efficacy endpoint) at Week 12 for each of the brazikumab groups is not different from that for the placebo group. The alternative hypothesis for each of the brazikumab groups is different from that of the placebo group, i.e.:

For the high dose,

$$H_0: p_{high} = p_{placebo} \ vs. H_1: p_{high} \neq p_{placebo}$$

For the low dose,

$$H_0: p_{low} = p_{placebo} \ vs. H_1: p_{low} \neq p_{placebo}$$

These hypotheses will be tested in hierarchical order starting with the high dose.

All other testing will be performed with nominal p-values; no multiplicity adjustments will be performed and secondary/exploratory tests should not be interpreted in terms of statistical significance.

3.3.5 Handling of Protocol Deviations in Study Analysis

A per-protocol (PP) analysis is not planned for this study. Important protocol deviations (IPD) will be listed and tabulated in the Clinical Study Report (CSR). Important protocol deviations are a subset of protocol deviations that may significantly impact the completeness, accuracy, or reliability of the study data or that may significantly affect a participant's rights, safety, or well-being.

IPDs will be documented prior to unblinding the study data for the Induction Period, and will be reported for the Induction Period at the Week 12 interim (primary) analysis. The final list of IPDs for the maintenance treatment Period will be reported at the Week 52 final analysis.

The protocol deviations, definitions, process for identification, and assessment are detailed in a separate protocol deviation plan.

4 STATISTICAL ANALYSIS

4.1 Study Population

The study population covers participant disposition, analysis sets, protocol deviations, demographics, baseline characteristics, medical history, prior and concomitant medications, and study intervention compliance.

4.1.1 Participant Disposition and Completion Status

4.1.1.1 Definitions and Derivations

Enrolled

A participant is enrolled if they, or their legally acceptable representative, sign the informed consent form (ICF).

Screen failure

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomly assigned to study intervention.

Randomized

Randomized participants are defined as those who undergo randomization and receive a randomization number.

Study discontinuation

A study discontinuation will occur if a participant who signs the ICF and is dosed ceases participation in the study, regardless of circumstances, before the completion of the protocol-defined study procedures.

Discontinuation of study intervention

A study intervention discontinuation will occur if a participant who has taken at least one dose of the study intervention is discontinued from the study intervention before the completion of protocol-defined study procedures.

Study intervention completion

A study intervention completer for Induction Period and/or Maintenance Period is defined as follows:

- If the participant has received the first dose of IV brazikumab and the dose at Week 8, the participant is considered to be a study intervention completer for the Induction Period.
- If the participant has received the first dose of SC brazikumab, and the dose at Week 48, and has not missed more than one dose, the participant is considered to be a study intervention completer for the Maintenance Period.

Study completion

Study completion for a participant means that the protocol-defined study procedures related to the study are completed.

4.1.1.2 Presentation

Listings and summaries will be presented for the Screened Analysis Set. The listings will include but not be limited to disposition event, time of event, and exposure. The summary will include the number and percentage of participants in the Screened Analysis Set who:

- Were screened
- Failed screening, and reasons
- Were randomized
- Randomized, did not receive study intervention
- Started Induction Period treatment
- Completed Induction Period treatment
- Rolled over into the OLE study during the Induction Period
- Discontinued IV treatment prior to Week 12, and reasons
- Discontinued SC treatment on Week 12, and reasons
- Withdrew from study during the Induction Period, and reasons
- Received Week 12 ileocolonoscopy
- Started Maintenance Period treatment
- Completed Maintenance Period treatment
- Rolled over into the OLE study during the Maintenance Period
- Discontinued Maintenance Period treatment, and reasons
- Completed Week 52
- Ongoing in the Maintenance Period (IA only)
- Withdrew from the study prior to Week 52, and reasons
- Started the Safety follow-up period
- Completed the Safety follow-up
- Ongoing in the Safety follow-up period (IA only)
- Discontinued during the Safety follow-up period, and reasons

Participant recruitment will be summarized overall and by brazikumab total and treatment group by country and study site. Participants randomized prior to CSP Amendment 4 will be included. The stratification factors (status of prior biologic use and current CS use [yes or

no] at randomization) will be summarized according to the IXRS vs. according to the eCRF, by treatment group and brazikumab total.

Percentages are based on the number of participants who started treatment in the induction period or in the maintenance period.

All participants who prematurely discontinue will be listed by period, i.e., the Induction Period, the Maintenance Period, and the Safety Follow-up Period with discontinuation reason for the Screened Analysis Set.

Summary of global/country situation A summary of disposition due to the global/country situation will be provided by treatment group, to include the number of participants who

- Completed treatment
- Discontinued treatment due to the global/country situation, and reasons
- Completed the study
- Withdrew from the study due to the global/country situation

In addition, global/country situation disruptions will be summarized for the FAS by treatment group, to include the number of participants with at least one disruption due to global/country situation:

- Visit impacted
- Study drug impacted
- Concomitant medication impacted
- Discontinued treatment due to global/country situation
- Withdrew from the study due to global/country situation

4.1.2 Analysis Sets

4.1.2.1 Definitions and Derivations

There are 5 analysis populations:

Screened Analysis Set

This analysis set includes all participants screened/enrolled for the study. Enrolled participants are defined as those who sign informed consent.

The Full Analysis Set (FAS)

This analysis set includes all participants who are randomized to a treatment group (as per the Intent-To-Treat (ITT) principle), excluding participants enrolled prior to CSP Amendment 4. The FAS will be used for all efficacy analyses unless otherwise indicated. Participants will be summarized according to the randomized study treatment group.

The Safety (SAF) Analysis Set

This includes all participants who receive ≥ 1 administration of study intervention including all participants enrolled prior to CSP Amendment 4. Participants who only received Humira® will only be included in the listings. Erroneously treated participants (e.g., those randomized to treatment A but actually given treatment B) are accounted for in the treatment group of the treatment they actually received. A participant who has on one or several occasions received active study intervention is classified as active and is accounted for in the active study intervention treatment group. A participant who has received both 720 mg and 1440 mg active dose regimens is classified as the higher active dose regimen. Participants will be summarized according to the actual study intervention received.

The PK Analysis Set

This includes Safety Set participants having at least 1 PK sample containing detectable brazikumab concentrations. Participants will be summarized according to the actual study treatment received. Participants who only received Humira® will only be included in the listings.

The ADA Evaluable Analysis Set

This includes Safety Set participants having a non-missing baseline ADA result and at least one non-missing post-baseline ADA result. Participants will be summarized according to the actual study treatment received. Participants who only received Humira® will only be included in the listings.

4.1.2.2 Presentation

The number of participants in each analysis set will be listed and summarized by treatment group, brazikumab total, and overall, including reasons excluded from the Full analysis set, Safety analysis set, PK analysis set, and ADA evaluable analysis set.

4.1.3 Protocol Deviations

4.1.3.1 Definitions and Derivations

The definition and derivation of IPDs is described in a separate Protocol Deviation plan. These IPDs will be reviewed and documented before the CDL for interim (primary) analysis when all participants have completed Week 12 or discontinued early, and again before final CDL for final analysis after all participants complete the last study visit (or discontinue early).

4.1.3.2 Presentation

The participants with IPDs will be listed. The number and percentages of participants with at least one IPD will be summarized overall and by treatment group.

IPDs summarized include:

- Violations of inclusion criteria
- Violations of exclusion criteria
- Discontinuation Criteria for study intervention are met but the participant was not withdrawn from study intervention
- Discontinuation Criteria for withdrawal from study are met but the participant was not withdrawn from the study
- Investigational product deviation
- Excluded medication taken
- Deviations related to study procedures
- Other important protocol deviations

4.1.4 Demographics

4.1.4.1 Definitions and Derivations

Demographic variables will include age, sex, region, country, race and ethnicity. Participant age is entered in years. Age will be further categorized in 3 groups (<40, 40-65, and >65).

Participants who reported multiple races will be counted in the category “multiple”.

4.1.4.2 Presentation

Demographic variables will be listed and summarized by treatment group, brazikumab total, and overall for the FAS and include:

- Age
- Age group [<40, 40-65, and >65]
- Sex
- Region [US and Non-US]
- Country
- Race
- Ethnicity

4.1.5 Baseline Characteristics

4.1.5.1 Definitions and Derivations

Baseline characteristics include:

- Height
- Weight
- BMI
- BMI group (<18.5, 18.5 – 24.9, 25-29.9, ≥ 30).

Baseline

Baseline and Change from Baseline are defined in Section 3.3.1.2.

4.1.5.2 Presentation

The baseline characteristics will be listed and summarized descriptively by treatment group, brazikumab total, and overall for the FAS Population.

4.1.6 Disease Characteristics

4.1.6.1 Definitions and Derivations

Disease characteristics to be collected at Baseline are listed in Table 4-1.

Table 4-1 Disease Characteristics to be Summarized at Baseline

Disease Characteristics at Baseline	• Disease duration of CD (years) • Current immunomodulator use • Baseline disease location • Fistula • Extra-intestinal manifestations • Prior use of JAK inhibitors • Prior biologic use (stratification factor) • Current CS use (stratification factor) • Number of prior biologics • Number of different mechanisms of action of biologics
Participant Responder status on prior Crohn's disease medications (Intolerance, Primary failure, Secondary failure)	• Anti-TNF (Infliximab, Adalimumab, Certolizumab) • Integrin receptor antagonist (Vedolizumab) • Anti-IL12/23 (Ustekinumab) • Other Biologic • Immunomodulator (Azathioprine, Methotrexate, 6-mercaptopurine) • Other immunomodulator
SES-CD (assessed during ileocolonoscopy)	• Total score • Score by baseline disease location subgroup (per subgroup of participants with ileum, right colon, transverse colon, left colon, and rectum-colonic disease)
CDAI (calculated from PRO, physician assessment and laboratory markers)	• 220-450 moderate to severe disease • >450 very severe disease

Disease duration will be calculated as the (Date of First Study Intervention–- Date of Onset of Crohn's disease)/365.25 yrs.

Immunomodulators will include either prior or concomitant use of azathioprine, methotrexate, 6-mercaptopurine, or other immunomodulator.

Baseline disease location will be grouped based on the following locations:

- Colon: anus, rectum, sigmoid, descending colon, transverse colon, ascending colon, cecum
- Not Colon or Ileum: gastric, duodenum, extraintestinal, jejunum
- Likely Ileum: Other Crohn's location

Any of these locations with the response = "Current" will be included.

Extra-intestinal manifestations (EIM) responses will be categorized as:

- Morning stiffness
- Arthropathy (peripheral arthropathy, axial arthropathy)
- Iritis/uveitis
- Skin problems (erythema nodosum, pyoderma gangrenosum, aphthous stomatitis/aphthous ulcers, anal fissures, anal abscess, non-anal abscess, fistula)
- Other

JAK inhibitors include prior use of tofacitinib or upadacitinib.

Prior biologic use at randomization: any medication identified as biologic and administered prior to randomization. See for examples of biologics.

Inadequate responders on prior biologics at baseline: For each prior biologic used, 'inadequate responder' will be defined as "Intolerance", "Primary failure", or "Secondary failure".

Number of different mechanisms of action of prior biologics: for each of the prior biologics identified as CD medication and are defined as inadequate responders as defined above, will be categorized as anti-TNF, integrin receptor antagonists, or IL-12/23 inhibitor per Table 4-2. The number of different mechanisms will be 0-3.

Table 4-2 Drug Names and ATC Codes

Category	*Drug name*	*ATC5 code*	*ATC4 code*
Immunomodulators	Azathioprine	L04AX01	L04AX
	6-mercaptopurine	L01BB02	L01BB
	Methotrexate	L01BA01/L04AX03	L01BA/L04AX
	Cyclosporine A	L04AD01	L04AD
	Tacrolimus	L04AD02	L04AD

	Other		
JAK inhibitors	Tofacitinib Upadacitinib	L04AA29 L04AA44	L04AA L04AA
BIOLOGICS			
Anti-TNF	Etanercept Infliximab Adalimumab Certolizumab Golimumab	L04AB01 L04AB02 L04AB04 L04AB05 L04AB06	L04AB L04AB L04AB L04AB L04AB
Integrin receptor antagonists	Vedolizumab Natalizumab	L04AA33 L04AA23	L04AA L04AA
IL-12 and -23 inhibitor	Ustekinumab	L04AC05	L04AC

CDAI baseline

CDAI is defined in Section 4.2.1.1 and CDAI Baseline in Section 3.3.1.2.

SES-CD

The SES-CD score is an endoscopic activity score used to assess the status and change of mucosal lesions. The score assesses 4 variables in up to 5 segments to yield its final result. The 5 segments assessed are:

- Rectum, defined as that portion distal to the rectosigmoid junction
- Left colon, including the sigmoid colon
- Transverse colon, defined as the segment between the hepatic and the splenic flexures
- Right colon, including the ileocecal valve, cecum, and ascending colon to the hepatic flexure
- Ileum

Each segment is assessed for four qualities: size of ulcers, ulcerated surface, affected surface, and presence of narrowing. The total score is the sum of these 20 components (see Appendix 7.3 for more details).

4.1.6.2 Presentation

The disease characteristics will be listed and summarized descriptively by treatment group, brazikumab total, and overall for the FAS.

4.1.7 Medical History and Concomitant Disease

4.1.7.1 Definitions and Derivations

Crohn's disease Medical History

Participants Crohn's disease medical history will be collected, including surgical history, chest x-ray, family history of colorectal cancer, tuberculosis history, and vaccination history.

General Medical and Surgical History

Verbatim terms on the eCRF form Medical and Surgical History/Physical Findings will be mapped to system organ class (SOC) and preferred term (PT). For each Finding, the item will be concurrent if it is ongoing at the time of enrollment.

Participants' medical and surgical histories and concurrent procedures will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), version 24.0 or newer.

4.1.7.2 Presentation

Crohn's disease medical history as described in Section 4.1.7.1 will be listed and summarized by treatment group, and overall for the FAS. Chest x-ray, vaccination history, family history of colorectal cancer, and tuberculosis history will be presented in a listing.

General medical and surgical histories will be summarized applying MedDRA terms. Each SOC and PT will be summarized by treatment group, brazikumab total, and overall for the FAS. Concurrent procedures will also be presented in a listing.

4.1.8 Prior and Concomitant Medications

4.1.8.1 Definitions and Derivations

The World Health Organization (WHO) Drug Dictionary Enhanced, March 2022 or newer, will be used to classify prior and concomitant medications by therapeutic class and drug name.

Prior medication

Prior medication is defined as any medication taken before the first dose of the study intervention.

Concomitant medication

Any medication taken by the participant at any time between the date of the first dose (including the date of the first dose) of IP up to 4 weeks after the last dose of IP, inclusive, is considered concomitant medication. Any medications taken by the participant prior to the first dose date of IP that continues after the first dose date of IP are considered both prior and concomitant medications. Any medications taken prior to the first dose date of IP whose dose has changed after the date of first dose of IP will be captured with a separate data entry with

the date of dose change, which will be used to determine if the medication is prior or concomitant.

Any medications started later than 4 weeks after the last dose of study intervention date will not be presented in the summary tables but will be included in the participant data listings.

Permitted Crohn's disease medication

Permitted Crohn's disease medications should be recorded as concomitant medications. See CSP Section 6.6.2 and Table 4-4 for details of permitted Crohn's disease medications.

Prohibited Crohn's disease intervention

Prohibited Crohn's disease intervention should not be taken during the course of the study. The interventions listed below are considered exclusionary and are not permitted through Week 52 of the study. In most cases, if prohibited interventions are taken, the study intervention must be discontinued. Use of prohibited intervention after randomization is considered an intercurrent event for all endpoints analyzed per the primary estimand (see Section 4.2.1.1).

Table 4-3 Prohibited medications

Medication
Biologic agents
Adalimumab
Certolizumab pegol
Infliximab
Golimumab
Vedolizumab
Ustekinumab
Risankizumab
Briakinumab
Mirikizumab
Guselkumab
Tildrakizumab
Natalizumab
Other biologic
Immunomodulator
Cyclosporine
Mycophenolate mofetil
Ozanimod (Zeposia)
Sirolimus (rapamycin)
Thalidomide
Tacrolimus (FK-506)
Tofacitinib
IV or intramuscular steroids
Others

Bacille Calmette-Guerin vaccination
Any live vaccine
Known or suspected history of chronic use of NSAIDs (not applicable to daily aspirin use up to 325 mg/day)
Known or suspected history of chronic use of opiates, drug, or alcohol abuse
Known or suspected abuse of marijuana, as judged by the investigator
Fecal microbiota transplantation
Chinese herbal therapies

Rescue treatment

Permitted rescue medications are allowed as per investigator judgment. Participants will continue to receive study intervention unless a prohibited medication is used.

Any new concomitant medication, or any increase in dose of a Baseline medication required to treat new or unresolved CD symptoms, except for antidiarrheal medications, will be considered rescue treatment. Any **new** initiation (or dose increase) of medications or interventions listed below will be considered rescue treatment:

Table 4-4 Permitted Rescue Medications

Medication	***ATC5 code***
5-aminosalicylates	A07EC02
Corticosteroids Parenteral, oral, or rectal CS Oral CS doses above Day 1 dose	 L04AD01 L04AD01
Immunomodulators Azathioprine 6-mercaptopurine Methotrexate	 L04AX01 L01BB02 L01BA01/L04AX03
IV immunoglobulin	J06BA02

CS doses up to the levels taken on Day 1 are not considered rescue treatment. Other investigational products are not permitted to be used as rescue treatment. Use of rescue medication after randomization is considered an intercurrent event for all endpoints analyzed per the primary estimand (see Section 4.2.1.1).

The criteria for use of rescue treatment are found in the CSP Section 6.6.4.

4.1.8.2 Presentation

Medications will be summarized for the FAS by the number and percentage of participants in each treatment group, brazikumab total, and overall receiving each medication within WHO drug class and generic drug name. If a participant took a specific medication multiple times or took multiple medications within a specific therapeutic class, that participant would be counted only once for the coded drug name or therapeutic class. Any prior and concomitant medications will be included in listings, including corticosteroid medications.

The following summaries will be provided:

- Prior medication use
- Prior corticosteroid use for Crohn's disease
- Prior immunomodulator use (for Crohn's disease and for other indication than Crohn's disease)
- Concomitant rescue medication use for Crohn's disease
- Permitted concomitant medication use
- Concomitant corticosteroid use (for Crohn's disease and for other indication than Crohn's disease)
- Concomitant immunomodulator use (for Crohn's disease and for other indication than Crohn's disease)
- Prohibited concomitant medication use

4.1.9 Study Intervention Compliance

4.1.9.1 Definitions and Derivations

Participants are dosed at the site and will receive study intervention directly from qualified blinded staff under medical supervision. The date and time of dose administered in the clinic will be recorded in the source documents and recorded in the CRF. Study intervention compliance will measure the number of doses received and will be calculated as the total number of doses taken during that period divided by the number of expected doses, multiplied by 100.

4.1.9.2 Presentation

Study intervention compliance will be listed and summarized descriptively for the Induction Period, the Maintenance Period, and Double-Blind Treatment Period by treatment group using the Safety analysis set. The number and percentage of doses received at Week 0, 4, 8, and during the Induction Period will be summarized by treatment group using the Safety analysis set. The mean (SD), median, and range of the number of doses received will be presented for the Maintenance and Double-Blind Periods by treatment group. Administration of investigational product will be presented in a listing.

4.2 Endpoint Analyses

This section covers details related to the endpoint analyses, including primary, secondary, and exploratory endpoints as well as sensitivity and supportive analyses.

Whenever parametric models are described, the data will undergo appropriate tests of the normality assumptions and if necessary, data will be transformed or non-parametric substitutes will be used for analysis.

Table 4-5 Objectives and Endpoints

Statistical category	*Endpoint*	*Analysis Population*	*Intercurrent event strategy*	*Population level summary (analysis)*	*Details in section*
Objective 1: To compare the efficacy of brazikumab with that of placebo to achieve CDAI remission at Week 12					
Primary	CDAI remission at Week 12: CDAI score < 150	FAS	NRI before Week 12 (see Section 4.2.1.1)	Percentage of participants achieving the endpoint	4.2.1
Objective 2: To compare the efficacy of brazikumab with that of placebo to achieve endoscopic response, CDAI response, and clinical remission at Week 12.					
Key Secondary	Endoscopic response at Week 12: Minimum of 50% decrease from Baseline in SES-CD total score	FAS	NRI before Week 12 (see Section 4.2.1.1)	Same as primary	4.2.2
Secondary	Clinical remission at Week 12: Average daily LSF subscore ≤ 3 as assessed on the CDAI LSF item AND average daily AP subscore of ≤ 1 as assessed on the CDAI AP item	FAS	NRI before Week 12 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	CDAI response at Week 12: CDAI score of <150 points or CDAI reduction from Baseline of ≥ 100 points	FAS	NRI before Week 12 (see Section 4.2.1.1)	Same as primary	4.2.3
Objective 3: To compare the efficacy of brazikumab with that of placebo to achieve sustained CDAI remission, CDAI response, endoscopic response, and clinical remission at both Week 12 and Week 52					
Secondary	CDAI remission at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	CDAI response at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	Endoscopic response at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3

Statistical category	*Endpoint*	*Analysis Population*	*Intercurrent event strategy*	*Population level summary (analysis)*	*Details in section*
Secondary	Clinical remission at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Objective 4: To compare the efficacy of brazikumab with that of placebo in achieving CDAI remission, CDAI response, endoscopic response, SES-CD total score of 0-2, endoscopic remission, and clinical remission at Week 52					
Secondary	Endoscopic remission at Week 52: SES-CD total score of 0-2 OR SES-CD total score of ≤ 4 and at least 2 point reduction from Baseline with no subscore > 1	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	Clinical remission at Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	CDAI response at Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	CDAI remission at Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	Endoscopic response at Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Secondary	SES-CD total score of 0-2 at Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Objective 5: To compare the efficacy of brazikumab with that of placebo to achieve endoscopic response at Week 12 and endoscopic remission at Week 52					
Secondary	Endoscopic response at Week 12 and endoscopic remission at Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.3
Objective 6: To evaluate the PK and immunogenicity of brazikumab in participants with CD					
Secondary	Population PK model of serum concentrations of brazikumab and analysis for serum anti-brazikumab antibodies	PK, ADA		Descriptive statistics. PK model to be developed in the PK analysis plan and reported outside the CSR.	4.2.44.4, 4.5
Objective 7: To characterize the exposure-response relationships of brazikumab					
Secondary	Exposure-response model linking primary endpoint to metrics of model-predicted individual brazikumab exposures	PK		Analysis to be developed in the PK analysis plan	4.2.4
Objective 8: To establish the serum IL-22 concentration Baseline clinical cutoff for its value in predicting the efficacy of brazikumab					

Statistical category	*Endpoint*	*Analysis Population*	*Intercurrent event strategy*	*Population level summary (analysis)*	*Details in section*
Secondary	Exploration of relationship of Baseline serum IL-22 concentration with efficacy of brazikumab at Week 12, and establishment of the serum IL-22 concentration clinical cutoff to stratify participants in Stage 2 through CDAI remission and endoscopic response at Week 12	FAS	NRI before Week 12 (see Section 4.2.1.1)	Differential effect method	4.2.5
Objective 9: To evaluate the safety and tolerability of brazikumab in participants with CD					
Secondary	AEs, clinical laboratory values, vital signs, physical exams, ECGs	Safety		Summary tables	4.6
Objective 10: To compare the efficacy of brazikumab with that of placebo to achieve CS-free CDAI remission, CDAI response, endoscopic remission, endoscopic response, SES-CD total score of 0-2, and clinical remission at Week 52 for participants taking CS at Baseline					
Exploratory	CS-free CDAI remission at Week 52 for participants taking CS at Baseline	FAS for subset of participants taking CS at Baseline	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	CS-free CDAI response at Week 52 for participants taking CS at Baseline	FAS for subset of participants taking CS at Baseline	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	CS-free endoscopic remission at Week 52 for participants taking CS at Baseline	FAS for subset of participants taking CS at Baseline	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	CS-free clinical remission at Week 52 for participants taking CS at Baseline	FAS for subset of participants taking CS at Baseline	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6

Statistical category	*Endpoint*	*Analysis Population*	*Intercurrent event strategy*	*Population level summary (analysis)*	*Details in section*
Exploratory	CS-free endoscopic response at Week 52 for participants taking CS at Baseline	FAS for subset of participants taking CS at Baseline	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	CS-free SES-CD total score of 0-2 at Week 52 for participants taking CS at Baseline	FAS for subset of participants taking CS at Baseline	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Objective 11: To compare the efficacy of brazikumab with that of placebo to achieve sustained endoscopic remission and SES-CD total score of 0-2 at both Week 12 and Week 52					
Exploratory	Endoscopic remission at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	SES-CD total score of 0-2 at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Objective 12: To compare the efficacy of brazikumab with that of placebo to achieve primary symptom remission at Week 12					
Exploratory	Primary symptom remission at Week 12 For participants with Baseline LSF subscore of ≥ 5 and AP subscore < 2: Average daily LSF subscore of ≤ 3 AND no worsening of Baseline AP subscore as assessed on the CDAI OR For participants with Baseline AP subscore of ≥ 2 and LSF subscore < 5: Average daily AP subscore of ≤ 1 AND no worsening of Baseline LSF subscore as assessed on the CDAI	FAS	NRI before Week 12 (see Section 4.2.1.1)	Same as primary	4.2.6
Objective 13: To compare the efficacy of brazikumab with that of placebo to achieve sustained primary symptom remission at both Week 12 and Week 52					
Exploratory	Primary symptom remission at both Week 12 and Week 52	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Objective 14: To compare the efficacy of brazikumab with that of placebo to achieve clinical response at Week 12					

Statistical category	*Endpoint*	*Analysis Population*	*Intercurrent event strategy*	*Population level summary (analysis)*	*Details in section*
Exploratory	Clinical response at Week12, defined as minimum 25% reduction in LSF subscore or AP subscore from Baseline	FAS	NRI before Week 12 (see Section 4.2.1.1)	Same as primary	4.2.6
Objective 15: To compare the efficacy of brazikumab with that of placebo in changes from baseline in CDAI					
Exploratory	Change from Baseline at Weeks 2, 4, 8, 12, 24, 40, and 52 in CDAI score	FAS	Accounted for by the mixed model (assumes MAR)	LS mean and 95% CI from a mixed model for repeated measures, assuming unstructured covariance matrix (if converges - compound symmetry otherwise)	4.2.7
Objective 16: To evaluate the impact of brazikumab on the signs and symptoms of CD through Week 52					
Exploratory	Change from Baseline for all Visits through Week 52 in signs and symptom scores (eg, LSF, AP, urgency, fatigue) derived from the evening diary and bowel movement diary; and FACIT-F at Weeks 12, 24, 40, and 52	FAS	Accounted for by the mixed model (assumes MAR)	LS mean and 95% CI from a mixed model for repeated measures, assuming unstructured covariance matrix (if converges - compound symmetry otherwise)	4.2.8
Objective 17: To evaluate the impact of brazikumab on HRQoL at Week 12 through Week 52					
Exploratory	Change from Baseline at Weeks 12, 24, 40, and 52, respectively for IBDQ, and Weeks 12, 24, and 52, respectively for SF-36, and EQ-5D-5L	FAS	Accounted for by the mixed model (assumes MAR)	LS mean and 95% CI from a mixed model for repeated measures, assuming unstructured covariance matrix (if converges - compound symmetry otherwise)	4.2.9
Objective 18: To explore changes in signs and symptoms and HRQoL using a range of measures					
Exploratory	Subscale scores from BMD, IBDQ, and additional symptom NRS items	FAS	Accounted for by the mixed model (assuming MAR)	LS mean and 95% CI from a mixed model for repeated measures, assuming unstructured covariance matrix (if converges - compound symmetry otherwise)	4.2.8, 4.2.9
Objective 19: To explore binary endpoints from Week 2 through Week 52					
Exploratory	CDAI response for Weeks 2, 4, 8, 12, 24, 40, and 52, respectively	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6

Statistical category	*Endpoint*	*Analysis Population*	*Intercurrent event strategy*	*Population level summary (analysis)*	*Details in section*
Exploratory	CDAI remission for Weeks 2, 4, 8, 12, 24, 40, and 52, respectively	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	Clinical remission for Weeks 2, 4, 8, 12, 24, 40, and 52, respectively	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	Both CDAI remission and endoscopic response at Week 12	FAS	NRI before Week 12 (see Section 4.2.1.1)	Same as primary	4.2.6
Exploratory	Both clinical remission and endoscopic response at Week 12	FAS	NRI before Week 52 (see Section 4.2.1.1)	Same as primary	4.2.6
Objective 20: To explore transcriptional, histological, protein, and microbiome biomarkers in study participants and their relationship with study intervention, disease features, clinical outcomes, and patient characteristics					
Exploratory	Changes from Baseline in IL-22, CRP, and FCP over time	FAS	Accounted for by the mixed model (assuming MAR)	LS mean and 95% CI from a mixed model for repeated measures, assuming unstructured covariance matrix (if converges - compound symmetry otherwise)	4.3
Exploratory	• Serum, plasma, fecal, or gut tissue proteins • whole blood or gut tissue transcriptional changes • histological and microbiome assessments	FAS		To be detailed in a separate exploratory analysis plan	
All primary, secondary, and exploratory efficacy endpoints will be evaluated in the BM+ and BM- populations as well					

AE = adverse event; AP = abdominal pain; BM = biomarker serum IL-22 concentrations at/above (BM+) or below (BM-), a pre-established cutoff; BMD = Bowel Movement Diary; BSFS = Bristol Stool Form Scale; CD = Crohn's disease; CDAI = Crohn's Disease Activity Index; CI = confidence interval; CRP = C-Reactive Protein; CS = corticosteroid; ECG = electrocardiogram; EQ-5D-5L = 5-level European Quality of Life – 5 Dimensions; FACIT-F = Functional Assessment of Chronic Illness Therapy – Fatigue Scale; FCP = Fecal calprotectin; HRQoL = health-related quality of life; IBDQ = Inflammatory Bowel Disease Questionnaire; IL-22 = interleukin-22; LS = least squares; LSF = loose stool frequency; MAR= missing at random; NRI = non-responder imputation; NRS = Numerical Rating Scale; PK = pharmacokinetics; SES-CD = Simple Endoscopic Score for Crohn's Disease; SF-36 = Short-Form 36 Health Survey.

4.2.1 Primary Endpoint

4.2.1.1 Definition

The primary study objective is to compare the treatment group level response probability (measured as percentage of participants with CDAI remission) between brazikumab groups versus placebo group in the FAS Population. The four components of the estimand are defined as follows:

A. The target population is defined through the inclusion/exclusion criteria for the study to reflect the moderately to severely active Crohn's disease patient population under investigation.
B. The variable is the same as the primary efficacy parameter, which is binary, indicating a successful response if there is a CDAI remission at Week 12
C. Study intercurrent events are captured in the variable definition as a participant who
   - discontinues the study intervention prematurely before Week 12
   - takes rescue treatment (specified in Section 4.1.8.1) before Week 12
   - uses prohibited medication (specified in in Section 4.1.8.1) before Week 12

   Participants who experience any of these intercurrent events are considered as being unsuccessfully treated (Composite Strategy – ICH E9 [R1]) and will be imputed as nonresponders (NRI).
D. The population-level summary measure is the percentage of participants with CDAI remission.

The estimand defined above is considered the primary estimand. In this case, the clinical question is the treatment effect on CDAI remission had no additional medication been available. CDAI remission values after rescue therapy or prohibited medications are not directly relevant since those values also reflect the impact of that additional medication and are considered as failure.

The null hypothesis is presented in Section 3.3.4.

4.2.1.2 Derivations

A participant will have achieved **CDAI remission** if their CDAI score is < 150.

The CDAI score is captured on the eCRF derived independently by RAVE, and the investigator. However, the CDAI score derived from the raw data of the individual components will be used in the statistical analyses.

The CDAI score is calculated by summing the weighted scores for the items in Table 4-6. The CDAI scores range approximately from 0 to 600, with higher scores indicating greater disease activity. Scores of < 150, 150 to 219, and 220 to 450 represent remission, mild disease,

and moderate to severe disease, respectively; whereas scores of > 450 represent very severe disease (Buxton et al, 2007).

To be noted that the CDAI PRO components (LSF, AP, and General Well-being) will be calculated by summing the individual daily scores over 7 days (see below for missing data conventions). This is in contrast with the average daily LSF and AP, which are calculated by taking the mean of the entries over the 7 days (see Appendix 7.2).

Table 4-6 Items Included in CDAI and Their Weights

Item	*Timing of assessment*	*×Weight*	Total
Total number of liquid or very soft stools over past week (LSF)	7 days prior*	× 2	X_1
Total abdominal pain (AP) score (rating: 0-3) over past week (range: 0-21)	7 days prior*	× 5	X_2
Total general well-being score (rating: 0-4) over past week (range: 0-28)	7 days prior*	× 7	X_3
Sum of presence of following clinical signs over past week: 1. Arthritis/arthralgia (1=yes, 0=no) 2. Iritis/ uveitis (1=yes, 0=no) 3. Erythema nodosum/Pyoderma gangrenosum/aphthous stomatitis (1=yes, 0=no) 4. Anal fissure, fistula or abscess (1=yes, 0=no) 5. Other fistula (1=yes, 0=no) 6. Fever > 37.8 C During Past Week (1=yes, 0=no)	 Assessed At Visit Assessed At Visit Assessed At Visit Assessed At Visit Assessed At Visit Assessed At Visit 7 days prior	× 20	X_4
Antidiarrheal use (e.g., diphenoxylate hydrochloride) (0=none, 1=yes)	Assessed At Visit	× 30	X_5
Abdominal mass (none=0, equivocal=2, present=5)	Assessed At Visit	× 10	X_6
Deviation of hematocrit levels (minimum value=0) 47 – hematocrit (males) 42 – hematocrit (females)	Lab collected at Visit	× 6	X_7
Weight ratio 100x(1-[Current body weight/standard weight]) Minimum = -10 for overweight participant Maximum = 10 for underweight participant (if value<-10, enter -10, if value >10, enter 10)	Assessed At Visit	× 1	X_8
		CDAI score	$\sum_{i=1}^{8} X_i$

Standard body weight (IBW) (men) = 50 kg + 2.3 kg x (height in cm / 2.54 – 60)

Standard body weight (IBW) (women) = 45.5 kg + 2.3 kg x (height in cm / 2.54 – 60)

Note: this formula is only an approximation and is generally only applicable for people 60 inches (152.4 cm) tall or greater.

* If the CDAI calculation is done on the same day as an ileocolonoscopy (e.g., Weeks 12, 52, E/T, and unscheduled visits), then the PRO elements will be collected during the 7 days prior to bowel prep.

Convention for calculations of CDAI scores:

1. Identify the CDAI calculation date as the completion date of the physician reported CDAI components using the Visit Window conventions in Section 3.3.2.
2. Calculate the 3 eDiary subscores (liquid/soft stool frequency, abdominal pain and general well-being) as follows:
   a. Select the diary data
      1) When no ileocolonoscopy is performed, select entries from 7 days prior to the CDAI calculation date identified in (1).
      2) When an ileocolonoscopy is to be performed on the same day as the CDAI assessment, select entries from 7 days prior to the initiation of bowel prep
   b. If 4 or more days of diary are non-missing, the subscore is calculated as the (average of non-missing diary x 7), multiplying the factor appropriate for the given subscore then rounding to the nearest integer
3. Calculate Extra-intestinal manifestations of Crohn's Disease Subscore, usage of Lomotil/Imodium/opiates for diarrhea and abdominal mass.
   a. The EIM subscore will be the sum of the six items as listed in Table 4-6
   b. Assign 1 point for use of diarrhea treatment
   c. Abdominal mass is scored as none=0, equivocal=2, present=5
4. Calculate Hematocrit subscore as follows:
   a. Identify the Hematocrit (%) results using the visit window defined in Section 3.3.2. Select the value closest to the CDAI calculation date in (1). If two hematocrits are in the visit window, follow the rules in Section 3.3.2.4.
   b. If a hematocrit is missing, use the previous Week's value, if available.
   c. To calculate the corresponding subscore subtract 47 from males and 42 for females, multiply by a factor of 6 and round to the nearest integer. If the hematocrit subscore is < 0, set it to 0.
5. Calculate Body Weight subscore as follows:
   a. Identify the body weight result reported on the CDAI calculation date.
   b. Identify the standard weight based on the participant's gender and baseline height (cm).
      1) Standard body weight (IBW)(men) = 50 kg + 2.3 kg × ((height in cm)/2.54 – 60)
      2) Standard body weight (IBW)(women) = 45.5 kg + 2.3 kg × ((height in cm)/2.54 – 60), or
   c. Calculate the subscore as (1-(Body weight/Standard weight)) x 100. If the subscore is below -10, set to -10. If the subscore is greater than 10, set to 10.

6. Calculate total score as the weighted sum of the 8 subscores only if none of the subscores are missing. Otherwise, the total score is set to missing for the visit. In addition, the following set of rules will be applied:

If any of the subscores cannot be calculated using the guidelines above, then rules explained in Section 4.2.1.3 will apply.

Any total CDAI score < 0 will be considered as 0 in the analysis.

4.2.1.3 Handling of Dropouts and Missing Data

NRI for intercurrent events will be applied prior to missing data handling rules.

Missing total CDAI score due to PRO components missing

If CDAI is missing due to missing PRO components, the following algorithm for handling PRO components missing data will be applied in order to obtain a CDAI score: in the event that there are not at least 4 days (consecutive or non-consecutive) of data needed to generate the weekly score, data from the week prior to the 7-day scoring window, if available, will be used to generate the weekly score. For example, if the period of interest is Day -7 to Day -1, and there were fewer than four days data complete, then consider Day -8 to Day -1. If there were still fewer than 4 days then consider Day -9 to Day -1, and so on until 4 days were observed.
This principle can only be applied within the 2 week dose delay threshold (as per CSP section 6.1.2) and without overlap. If the CDAI score cannot be calculated, then NRI rules apply.

Missing total CDAI score due to other than PRO components missing

If CDAI total score is missing due to reasons other than PRO components missing, the component calculated at the most recent prior visit (Unscheduled Visit or Week 8 Visit) can be imputed using last observation carried forward (LOCF) if only a single visit has missing CDAI (i.e. a missing observation can only be carried forward once). If the CDAI score still cannot be calculated, then NRI rules apply.

Missing total CDAI score due to both PRO and non-PRO components missing

If both PRO and non-PRO components are missing, the CDAI score calculated at the most recent prior visit (Unscheduled Visit or Week 8 Visit) can be imputed using last observation carried forward (LOCF) if only a single visit has missing CDAI (i.e. a missing observation can only be carried forward once). If the CDAI score still cannot be calculated, then NRI rules apply.

See Section 4.2.1.2 for more details on the calculation of the total CDAI score.

4.2.1.4 Primary Analysis of Primary Endpoint

The null hypothesis is described in Section 3.3.4. Statistical comparisons will be explored for each brazikumab dose group versus placebo.

For the primary efficacy endpoint, the observed response proportions will be calculated by treatment group in the FAS. The efficacy endpoint will be analyzed using a Cochran-Mantel-Haenszel (CMH) test controlling for the 2 randomization stratification factors (status of prior biologic use and current CS use [yes or no] at randomization). If the test statistic cannot be calculated (e.g. there is a stratum with no records), then current CS use strata will be dropped. Each dose will be tested hierarchically (high dose then low dose, only if the null hypothesis is rejected for the high dose).

The CMH test analyzes the association between treatment (brazikumab and placebo) and a binary outcome (CDAI remission) while taking into account the stratification factors.

There are n_{ij} participants in each stratum, where i is the stratum and j is the treatment group. The number of participants achieving CDAI remission is x_{ij} and the proportion of participants achieving CDAI remission is $p_{ij} = x_{ij} / n_{ij}$.

For each stratum, the difference in proportion of participants achieving CDAI remission is calculated as $d_i = p_{iB} - p_{iP}$, where B and P denote the different treatment groups (brazikumab and placebo, respectively).

Weights for each stratum are calculated as

$$w_i = \frac{n_{iB} * n_{iP}}{n_{iB} + n_{iP}}$$

Then the weighted difference (WD) is calculated as

$$WD = \frac{\sum w_i d_i}{\sum w_i}$$

The standard error (SE) of the weighted difference under the null hypothesis is given by

$$SE = \sqrt{\frac{\sum[w_i^2 Var(d_i)]}{(\sum w_i)^2}}$$

Where

$$Var(d_i) = \frac{p_i(1-p_i)n_i}{w_i(n_i-1)}$$

And

$$p_i = \frac{x_i}{n_i} = \frac{x_{iB} + x_{iP}}{n_{iB} + n_{iP}}$$

To derive the confidence interval (CI) for the weighted difference in proportions, a correction will be applied to the variance to make it more accurate. For each strata,

$$Var(d_i) = \frac{p^*_{iB}(1-p^*_{iB})}{n_{iB}} + \frac{p^*_{iP}(1-p^*_{iP})}{n_{iP}}, where\ p^*_{ij} = \frac{x_{ij}+2}{n_{ij}+4}$$

The 95% CI is $WD \pm z_{0.975} * SE$.

If the resulting upper or lower limit is >100% or < -100%, it will be set to 100% or -100%, respectively.

The value of the test statistic is *WD/SE,* and the p-value from the two-sided test of no difference in treatment groups is calculated as

$$p = 2\,(1 - Prob(|{}^{WD}/_{SE}|))$$

Where $Prob(|{}^{WD}/_{SE}|)$ is the distribution function of the standard normal distribution.

4.2.1.4.1 Presentation

The following analysis summary statistics will be presented for the primary efficacy endpoint on the FAS population by treatment and overall:

- Responder counts and response rate for each treatment group
- Estimated treatment effect (i.e., difference in response rate for active versus placebo), and corresponding 95% CI, and 2-sided p-value
- Number and percent of participants who are NRI due to:
    - Intercurrent event
        - Discontinues study intervention prematurely
        - Takes rescue medication
        - Uses prohibited medication
    - Missing data

NRI will be assessed for intercurrent events, then for missing data. If a participant experiences more than one category of intercurrent event, that participant will be counted in each category.

A graph presenting CDAI remission over time will be presented. In addition, a listing will be presented which includes, but is not limited to, whether the participant is a responder and which (if any) intercurrent events occurred.

4.2.1.5 Sensitivity Analyses of the Primary Endpoint

To assess the impact of baseline BM on the primary efficacy endpoint, a logistic regression model will be performed using treatment group, and the 2 stratification variables (status of prior biologic use and current CS use [yes or no] at randomization) as factors and baseline serum IL-22 value as a continuous covariate. The dependent variable for this sensitivity analysis is for the same primary estimand (CDAI remission), and the Delta method will be used to estimate the rate differences and CIs (Guo et al, 2012).

A second sensitivity analysis will be performed with baseline serum IL-22 categorized into BM+ and BM- as identified in Section 4.2.5. If the model fails to converge with three factors, then each factor will be analyzed individually as described in Section 4.2.1.7.

4.2.1.6 Supplementary Analyses of the Primary Endpoint

Treatment policy strategy

Per the treatment policy strategy, the actual value for CDAI remission is used regardless of whether or not an intercurrent event occurs, i.e., the values of the variable of interest are used whether or not the following intercurrent events occur:

- Rescue therapy before Week 12
- Any prohibited treatments before Week 12

In this case, NRI will be used only for the following intercurrent event

- Discontinue from study intervention before Week 12

The clinical question is the treatment effect on CDAI remission regardless of which other therapies are to be used before CDAI remission is experienced. Actual CDAI remission values after the use of additional treatment will be relevant.

The statistical analysis is the same as that for the primary efficacy analysis described in Section 4.2.1.4.

4.2.1.7 Subgroup Analyses

The same analysis as described in Section 4.2.1.4 will be performed for the following subgroups:

- Baseline serum IL-22 (High/BM+, Low/BM-), using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5)
- Prior biological use (naïve vs prior use)
- Current CS use at randomization (Y, N)
- Baseline CRP High (CRP+, ≥ 5 mg/L) vs. CRP Low (CRP-, <5 mg/L)
- Baseline FCP High (FCP+, ≥ 250 µg/g) vs. FCP Low (FCP-, < 250 µg/g)

- Region/Country (US, non-US)
- Race (White, non-White)
- Gender (Male, Female)
- Age group (<40, 40 to 65, >65)
- Baseline serum IL-22 (High/BM+, Low/BM-) and prior biological use (naïve vs prior use)

The analysis will be conducted separately on each subgroup. In case the test statistic cannot be calculated (e.g. there is a stratum with no records) for any of these analyses, the current CS use strata will be dropped. If the test statistic cannot be calculated in spite of this, the prior biologic use strata will be dropped and an unstratified CMH test will be used. Serum IL-22 subgroup classification will be determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5). For testing of "prior biological use" and "Current CS use at randomization" as subgroups, the corresponding strata will be dropped. A forest plot will be presented.

4.2.2 Key Secondary Endpoint

4.2.2.1 Definition

The key secondary objective is to compare the treatment group level response probability (measured as percentage of participants with endoscopic response) between brazikumab groups versus placebo in the FAS Population. The four components of the estimand are defined as follows:

A. The target population is defined through the inclusion/exclusion criteria for the study to reflect the moderately to severely active Crohn's disease patient population under investigation.
B. The variable is the key secondary efficacy parameter, which is a binary, indicating a successful response if there is an endoscopic response at Week 12
C. Study intercurrent events are captured in the variable definition as a participant who
    - discontinues the study intervention prematurely before Week 12
    - takes rescue treatment (specified in Section 4.1.8.1) before Week 12
    - uses prohibited medication (specified in Section 4.1.8.1) before Week 12

    Participants who experience any of these intercurrent events are considered as being unsuccessfully treated (Composite Strategy – ICH E9 [R1]) and will be imputed as nonresponders (NRI).
D. The population-level summary measure is the percentage of participants with endoscopic response.

4.2.2.2 Derivations

Endoscopic response will be derived using the following formula:

If (baseline total SES-CD–- total SES-CD at week 12)/baseline total SES-CD > 0.5 endoscopic response will be set to 1 (endoscopic responder), otherwise it will set to 0 (endoscopic non-responder). SES-CD total score is defined in Appendix 7.3. Baseline is defined in Section 3.3.1.2.

4.2.2.3 Handling of Dropouts and Missing Data

The endoscopy data received by the Vendor, based on Central Reader adjudication, are not going to be imputed. The vendor applies the following handling of missing data: if a post-baseline SES-CD segment is missing, the subscore for that SES-CD segment will be imputed to zero. The imputed segment scores will then be used for the SES-CD total score calculation.

If SES-CD total score is missing, then the endoscopic response will be imputed as a nonresponder.

4.2.2.4 Primary Analysis of Key Secondary Endpoint

Analysis will be performed on the FAS for the key secondary endpoint as described in Section 4.2.1.4. CMH will be stratified by prior biological use status and current corticosteroid use at randomization.

4.2.2.4.1 Presentation

Results of the analysis will be listed and summarized as described in Section 4.2.1.4.1.

4.2.2.5 Sensitivity Analyses of the Key Secondary Endpoint

Sensitivity analyses will be performed as for the Primary Endpoint described in Section 4.2.1.5.

4.2.2.6 Supplementary Analyses of the Key Secondary Endpoint

A supplementary analysis will be performed on the key secondary endpoint using treatment policy strategy similar to the supplementary analysis for the primary endpoint, as described in Section 4.2.1.6.

4.2.2.7 Subgroup Analyses

The subgroup analyses for endoscopic response will be the same as those described in Section 4.2.1.7.

4.2.3 Secondary Endpoints: Binary Endpoints

The secondary binary endpoints include:

- Clinical remission at Week 12
- CDAI response at Week 12

- CDAI remission at both Week 12 and Week 52
- CDAI response at both Week 12 and Week 52
- Endoscopic response at both Week 12 and Week 52
- Clinical remission at both Week 12 and Week 52
- Endoscopic remission at Week 52
- Clinical remission at Week 52
- CDAI response at Week 52
- CDAI remission at Week 52
- Endoscopic response at Week 52
- SES-CD total score of 0-2 at Week 52
- Endoscopic response at Week 12 and endoscopic remission at Week 52

4.2.3.1 Definition

CDAI remission is defined in Section 4.2.1.1.

Endoscopic response is defined in Section 4.2.2.1.

Clinical remission, CDAI response, and **endoscopic remission** are defined within the derivation section below.

Participants who experience any of the intercurrent events defined in Section 4.2.1.1 prior to a planned time point are considered as being unsuccessfully treated (Composite Strategy – ICH E9 [R1]) for the corresponding endpoints and will be imputed as nonresponders (NRI). In addition to the intercurrent events defined in Section 4.2.1.1, the following intercurrent event is defined for time points after Week 12:

- Failure to taper steroids (further details in CSP Section 9.4.1)

A treatment policy strategy will be used to handle the intercurrent event of failure to taper steroids. That is, the values of the outcome will be used regardless if this intercurrent event occurs or not.

4.2.3.2 Derivation

The CDAI score and SES-CD total score are used to derive the endpoints in this section. CDAI score is described in Section 4.2.1.2 and SES-CD total score is described in Appendix 7.3.

For the following calculations, Baseline and percent Change from Baseline are defined in Section 3.3.1.2.

Clinical remission at Week 12 is defined as an average daily LSF subscore of ≤ 3 as assessed on the CDAI LSF item AND average daily AP subscore of ≤ 1 as assessed on the CDAI AP item at Week 12.

The CDAI LSF item is the average daily LSF subscore from the Evening Diary, based on the 7 days prior to initiation of bowel prep for the Week 12 or Week 52 ileocolonoscopy.

The CDAI AP item is the average daily abdominal pain score from the Evening Diary, based on the 7 days prior to initiation of bowel prep for the Week 12 or Week 52 ileocolonoscopy (Appendix 7.2).

CDAI response at Week 12 is calculated as a CDAI score of <150 points OR CDAI reduction from Baseline of ≥ 100 points at Week 12.

CDAI remission at both Week 12 and Week 52 is derived as CDAI total score <150 at both Week 12 and Week 52. A participant must be a responder for CDAI remission at both Week 12 and Week 52 to be a responder for this measure.

CDAI response at both Week 12 and Week 52: calculation at Week 12 is as above. CDAI response at Week 52 is similarly calculated. A participant must be a responder for CDAI response at both Week 12 and Week 52 to be a responder for this measure.

Clinical remission at both Week 12 and Week 52 is defined as an average daily LSF subscore of ≤ 3 as assessed on the CDAI LSF item AND average daily AP subscore of ≤ 1 as assessed on the CDAI AP item at Week 12 and Week 52. A participant must be a responder for clinical remission at both Week 12 and Week 52 to be a responder for this measure.

Endoscopic response at both Week 12 and Week 52 is defined as a minimum of 50% decrease from Baseline in SES-CD total score at both Week 12 and Week 52. A participant must be a responder for endoscopic response at both Week 12 and Week 52 to be a responder for this measure.

Endoscopic remission at Week 52 is defined as SES-CD total score of 0-2, OR SES-CD total score of ≤ 4 and at least 2 point reduction from Baseline with no subscore >1, using data from the Week 52 ileocolonoscopy.

Clinical remission at Week 52 is defined as an average daily LSF subscore of ≤ 3 as assessed on the CDAI LSF item AND average daily AP subscore of ≤ 1 as assessed on the CDAI AP item at Week 52.

CDAI response at Week 52 is defined as CDAI score of <150 points or CDAI reduction from Baseline of ≥ 100 points at Week 52.

CDAI remission at Week 52 is defined as CDAI < 150 at Week 52.

Endoscopic response at Week 52 is defined as a minimum of 50% decrease from Baseline in SES-CD total score, using data from the Week 52 ileocolonoscopy.

SES-CD total score of 0-2 at Week 52 is assessed using data from the Week 52 ileocolonoscopy.

Endoscopic response at Week 12 and endoscopic remission at Week 52 A participant must be a responder for endoscopic response at Week 12 and endoscopic remission at Week 52 to be a responder for this measure.

4.2.3.3 Handling of Dropouts and Missing Data

Missing CDAI elements will be handled as described in Section 4.2.1.3. Following the principle of "LOCF just once", LOCF will be used only for the first missing value. That is, missing CDAI at Week 12 can be imputed from Week 8, and missing CDAI at Week 52 can be imputed from Week 40 (or a more recent unscheduled visit). If Week 12 or Week 52 CDAI cannot be imputed, then the corresponding endpoint will be imputed to nonresponder.

Similarly, handling of missing data for SES-CD based endpoints is described in Section 4.2.2.3. If SES-CD total score is missing, then the corresponding endpoint will be imputed to nonresponder.

4.2.3.4 Primary Analysis of Secondary Endpoints

Analysis for each binary endpoint will be performed on the FAS population as for the Primary Endpoint as described in Section 4.2.1.4.

4.2.3.4.1 Presentation

Results of the analyses will be presented as described in Section 4.2.1.4.1.

4.2.3.5 Additional Analyses of the Secondary Endpoints

Supplementary analyses will be performed on selected secondary endpoints using a Treatment Policy strategy similar to the supplementary analysis for the primary endpoint as described in Section 4.2.1.6. The following endpoints will be analyzed:

- Clinical remission at Week 12
- Clinical remission at Week 52
- Endoscopic remission at Week 12
- Endoscopic remission at Week 52
- CDAI remission at Week 52

4.2.3.6 Subgroup Analysis

The same analysis as described in Section 4.2.1.7 will be performed for the secondary binary endpoints by the Baseline serum IL-22 (High/BM+, Low/BM-), using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5).

4.2.4 Secondary Endpoint: Exposure-Response Model

An exposure-response model will link the primary endpoint to metrics of model-predicted individual brazikumab exposures. This will be described in the PK SAP.

4.2.5 Secondary Endpoint: IL-22 Cutoff Determination

4.2.5.1 Definition

IL-22 is a cytokine produced by immune cells at sites of inflammation. The objective is the exploration of relationship of Baseline serum IL-22 concentration with efficacy of brazikumab at Week 12, and establishment of the serum IL-22 concentration clinical cutoff to stratify participants in Stage 2. The defined cutoff will also be used to classify participants into BM+ and BM- groups in this Stage 1 study.

4.2.5.2 Derivations

Biomarker status (BM+ or BM-) will be defined using IL-22, and CDAI data. The IL-22 concentration cutoff will be selected as the value that maximizes the difference in CDAI remission at Week 12 (defined in Section 4.2.1.2) between brazikumab and placebo when comparing the BM+ and BM- subpopulations for the pooled doses of brazikumab vs placebo. The cutoff used for BM status classification in the rest of this SAP is the value that maximizes the differential effect of CDAI remission at Week 12. This cutoff will be algorithmically derived using normalized Z-scores across the population of all possible cutoffs using the differential effects method (See Appendix 7.9). The IL-22 cutoff to be used in Stage 2 will be selected with consideration to the totality of statistical evidence in conjunction with clinical, regulatory, and commercial relevance.

4.2.5.3 Handling of Dropouts and Missing Data

The IL-22 cutoff will be derived from CDAI remission (Section 4.2.1.2). Handling of missing CDAI data is described in Section 4.2.1.3, while handling of missing SES-CD data is described in Section 4.2.2.3.

The Baseline IL-22 value will be used to determine the cutoff value. If IL-22 is missing at Baseline, that observation will not be used in the analysis.

4.2.5.4 Primary Analysis of Secondary Endpoint

The differential effect method (see Appendix 7.9) will be applied separately for the following variables: CDAI remission, and endoscopic response at Week 12 for each of the brazikumab doses and for the pooled brazikumab doses versus placebo. All analyses will use the FAS. Prevalence of the BM+ for each scenario will be provided. The cutoff that will be used for all BM subgroup analyses, as reported in the CSR, will be that derived at Week 12 from CDAI remission using the pooled brazikumab doses versus placebo. All other cutoffs are derived for exploratory purposes only.

4.2.5.4.1 Presentation

Graphs describing the relationship between all available baseline IL-22 cutoff points vs the Z-score (expressed by $g(h)$ in Appendix 7.9) of CDAI remission, and endoscopic response at Week 12 will be produced separately for each of the brazikumab doses and for the pooled brazikumab doses versus placebo.

In addition, descriptive summaries with the proportions of responders for CDAI remission, and endoscopic response at Week 12 for each of the brazikumab doses and for the pooled brazikumab doses versus placebo by IL-22 cutoff level will be produced together with graphs of the proportion of responders as a function of IL-22 cutoff levels by treatment group. Prevalence of BM+ for the continuous range of cutoff points will be presented for the overall group in a graph.

4.2.6 Exploratory Endpoints: Binary Endpoints

The exploratory binary endpoints include:

- CS-free CDAI remission at Week 52 for participants taking CS at Baseline
- CS-free CDAI response at Week 52 for participants taking CS at Baseline
- CS-free endoscopic remission at Week 52 for participants taking CS at Baseline
- CS-free clinical remission at Week 52 for participants taking CS at Baseline
- CS-free endoscopic response at Week 52 for participants taking CS at Baseline
- CS-free SES-CD total score of 0-2 at Week 52 for participants taking CS at Baseline
- Endoscopic remission at both Week 12 and Week 52
- SES-CD total score of 0-2 at both Week 12 and Week 52
- Primary symptom remission at Week 12
- Primary symptom remission at both Week 12 and Week 52
- Clinical response at Week 12
- CDAI response at Weeks 2, 4, 8, 12, 24, 40, and 52, respectively
- CDAI remission at Weeks 2, 4, 8, 12, 24, 40, and 52, respectively
- Clinical remission at Weeks 2, 4, 8, 12, 24, 40, and 52, respectively
- Endoscopic remission at Weeks 12 and 52, respectively
- SES-CD total score 0-2 at Weeks 12 and 52, respectively
- Both CDAI remission and endoscopic response at Week 12
- Both clinical remission and endoscopic response at Week 12

4.2.6.1 Definition

CDAI score, CDAI LSF and AP subscores, and SES-CD total score are defined in Section 4.2.1.2, Appendix 7.2, and Appendix 7.3 respectively. The estimand framework is described

in Section 4.2.1.1 for those endpoints based on CDAI or its components, and in Section 4.2.2.1 for endpoints based on SES-CD.

CDAI remission is defined in Section 4.2.1.1.

Endoscopic response is defined in Section 4.2.2.1.

CDAI response, endoscopic remission, and **clinical remission** are described in Section 4.2.3.2.

Clinical response, and p**rimary symptom remission** are defined within the derivation section below.

Participants who experience any of the intercurrent events defined in Section 4.2.1.1 prior to a planned assessment are considered as being unsuccessfully treated (Composite Strategy – ICH E9 [R1]) for the corresponding endpoints and will be imputed as nonresponders (NRI). In addition to the intercurrent events defined in Section 4.2.1.1, the following intercurrent event is defined for time points after Week 12:

- Failure to taper steroids (further details in CSP Section 9.4.1)

Unless otherwise specified, a treatment policy strategy will be used to handle the intercurrent event of failure to taper steroids. That is, the values of the outcome will be used regardless if this intercurrent event occurs or not.

For the CS-free endpoints, a composite variable strategy will be used, where the intercurrent event of failure to taper steroids will be handled in a conservative manner: participants with any CS use during the last 12 weeks before the date of assessment are considered as being unsuccessfully treated for the corresponding endpoints and will be imputed as nonresponders (NRI).

4.2.6.2 Derivation

CS-free CDAI remission at Week 52 for participants taking CS at Baseline:

Participants who were taking CS at Baseline, did not take CS for the last 12 weeks before the date of assessment at Week 52 and achieved CDAI remission at Week 52.

CS-free CDAI response at Week 52 for participants taking CS at Baseline:

Participants who were taking CS at Baseline, did not take CS for the last 12 weeks before the date of assessment at Week 52 and achieved CDAI response at Week 52.

CS-free endoscopic remission at Week 52 for participants taking CS at Baseline:

Participants who were taking CS at Baseline, did not take CS for the last 12 weeks before the date of assessment at Week 52 and achieved endoscopic remission at Week 52.

CS-free clinical remission at Week 52 for participants taking CS at Baseline:

Participants who were taking CS at Baseline, did not take CS for the last 12 weeks before the date of assessment at Week 52 and achieved clinical remission at Week 52.

CS-free endoscopic response at Week 52 for participants taking CS at Baseline:

Participants who were taking CS at Baseline, did not take CS for the last 12 weeks before the date of assessment at Week 52 and achieved endoscopic response at Week 52.

CS-free SES-CD total score of 0-2 at Week 52 for participants taking CS at Baseline:

Participants who were taking CS at Baseline, did not take CS for the last 12 weeks before the date of assessment at Week 52 and achieved SES-CD total score of 0-2 at Week 52.

Endoscopic remission at both Week 12 and Week 52:

Participants who achieved endoscopic remission at both Week 12 and Week 52.

SES-CD total score of 0-2 at both Week 12 and Week 52:

Participants who achieved SES-CD total score of 0-2 at both Week 12 and Week 52.

Primary symptom remission at Week 12:

1. For participants with Baseline LSF subscore of ≥ 5 and AP subscore < 2: Average daily LSF subscore of ≤ 3 AND no worsening of AP subscore as assessed on the CDAI OR

2. For participants with Baseline AP subscore of ≥ 2 and LSF subscore of < 5: Average daily AP subscore of ≤ 1 AND no worsening of LSF subscore as assessed on the CDAI.

Only participants meeting the two initial conditions will be analyzed for this endpoint.

Primary symptom remission at both Week 12 and Week 52:

Participants who achieved primary symptom remission at both Week 12 and Week 52.

Clinical response at Week 12:

A minimum 25% reduction in CDAI LSF subscore or AP subscore from Baseline. LSF and AP subscores will be taken as the mean of the last 7 days prior to the CDAI assessment and prior to initiation of bowel prep for the Week 12 and Week 52 ileocolonoscopy.

Both CDAI remission and endoscopic response at Week 12:

Participants who achieved both CDAI remission and endoscopic response at Week 12.

Both clinical remission and endoscopic response at Week 12:

Participants who achieved both clinical remission and endoscopic response at Week 12.

4.2.6.3 Handling of Dropouts and Missing Data

The exploratory binary endpoints are based on the CDAI score, SES-CD total score, CDAI-LSF, and CDAI-AP, and will be handled similarly to the secondary binary endpoints as described in Section 4.2.3.3.

4.2.6.4 Primary Analysis of Exploratory Endpoints

The binary exploratory endpoints will be analyzed for the FAS using CMH as defined in Section 4.2.1.4.

4.2.6.4.1 Presentation

Results of the analysis will be listed and summarized as described in Section 4.2.1.4.1.

4.2.6.5 Subgroup Analysis

The same analysis as described in Section 4.2.1.7 will be performed for the exploratory endpoints by the Baseline serum IL-22 (High/BM+, Low/BM-), using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5).

4.2.7 Exploratory Endpoint: Change from Baseline in CDAI score

4.2.7.1 Definition

This endpoint is the Change from Baseline in CDAI for each scheduled CDAI timepoint (Weeks 2, 4, 8, 12, 24, 40, and 52). CDAI score is described in Appendix 7.1.

4.2.7.2 Derivations

Change from Baseline will be calculated as defined in Section 3.3.1.2.

4.2.7.3 Handling of Dropouts and Missing Data

Handling of missing CDAI components and total scores is described in Section 4.2.1.3. To be noted that a baseline value cannot be carried forward to a post baseline assessment. This implies that, for example, if Week 2 assessment is missing, the Week 2 assessment will be missing as the baseline assessment cannot be carried forward to a post baseline. If a CDAI assessment is missing for a particular Week, then the CDAI score will be set to missing for that Week.

Missing data will be handled in the mixed model. Mixed models for repeated measures (MMRM) does not employ formal imputation; rather, it uses all available data (including participants with partial data) to estimate the mean treatment effect in an unbiased manner.

4.2.7.4 Primary Analysis of Exploratory Endpoint

For participants i=1, …, I and repeated observations j=1,…,J, the MMRM model can be parameterized as

$$Y_i = X_i\beta + \varepsilon_i$$

Where

$\mathbf{Y}_i$ = a J_i dimensional vector of outcome measurements for the i^{th} participant,

$\boldsymbol{\beta}$ = a p dimensional vector containing the fixed effects

X_i = ($J_i \times p$) dimensional design matrices of known covariates

$\boldsymbol{\varepsilon_i}$ = J_i dimensional vector of residual components ($\boldsymbol{\varepsilon_i} \sim N(0, \Sigma_i)$)

Σ_i = a ($J_i \times J_i$) covariance matrix which depends on i only through its dimension J_i

The response of Change from Baseline in CDAI will be analyzed via a MMRM which will include fixed effects for baseline CDAI score as a continuous covariate and treatment group, time point, treatment group-by-time point interaction, and the stratification variables (status of prior biologic use and current CS use [yes or no] at randomization) as factors. The Kenward and Roger method will be used to calculate the denominator degrees of freedom for the test of fixed effects. All visits with a scheduled CDAI assessment will be considered for inclusion in the model. The MMRM will only include time points that contain at least 10 non-missing measurements per treatment group.

Using the FAS, MMRM will be used to generate the least squares (LS) mean and 95% CI for each Week's timepoint, assuming unstructured covariance matrix. If use of the covariance matrix does not converge, repeat using compound symmetry (then AR(1), Toeplitz, Heterogeneous Toeplitz, and Heterogeneous Compound Symmetric, in that order).

4.2.7.4.1 Presentation

The Change from Baseline at each Week will be presented using descriptive statistics for each score. Least square (LS) means, standard error, and 95% CIs will be presented by treatment group, as well as the LS mean difference comparisons between brazikumab groups and placebo, and the standard error and 95% CI of the difference will be presented at each Week. Additionally, summary graphs of LS mean CDAI change from baseline over time will be presented by treatment and by BM status.

4.2.7.5 Subgroup Analysis

A subgroup analysis of Change from Baseline in CDAI score will be performed for the subgroup based on Baseline serum IL-22 (High/BM+, Low/BM-), using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5). The model used is as specified in Section 4.2.7.4 with the addition of the following fixed effects: subgroup, treatment group-by-subgroup, time point-by-subgroup, and treatment group-by-subgroup-by time point as factors.

4.2.8 Exploratory Endpoint: Change in Signs and Symptoms (BMD, Evening Diary, FACIT-F)

4.2.8.1 Definition

The exploratory endpoints are Change from Baseline through Week 52 in signs and symptoms derived from the Bowel Movement Diary, Evening Diary, and Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F). Frequency and type of stool (stool, blood, mucus), stool consistency (using BSFS) and urgency are collected from the BMD. Worst abdominal pain, worst fatigue, worst tiredness, worst weakness, worst lack of energy, and worst joint pain, are collected from the Evening eDiary (Appendix 7.5). BMD and Evening Diary scores are evaluated at Weeks 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, and 52. For visits where no ileocolonoscopy is performed, averages will be generated within the 7 days prior to each Visit. For visits where ileocolonoscopy is performed, averages will be generated during the 7 days prior to bowel prep. The FACIT-F is assessed on-site during Weeks 12, 24, 40, and 52.

The FACIT-F is a 13-item instrument that measures fatigue in chronic illness patients. See Appendix 7.6.

BMD items

Average stool frequency is derived from the number of entries in the BMD, including any BM entries entered in the Evening Diary.

Stool type includes the responses Stool, Blood, and Mucus on question E04 (Table 7-4). For instance, for a given day, it is possible that stool =7, blood=1, and mucus=2 and these will be averaged separately.

Stool consistency will be derived from question E06 (Table 7-4).

Urgency will be derived from question E05 (Table 7-4).

Evening diary

The Evening Diary contains 6 numeric rating scale (NRS) items with responses from 0 (none) to 10 (as bad as I can imagine), in which the participant rates the symptom at its worst in the past 24 hours:

- Abdominal pain
- Fatigue
- Tiredness
- Weakness
- Lack of energy

- Joint pain

In addition, the following CDAI items related to signs and symptoms are measured:

- Loose stool frequency (number of liquid or very soft stools)
- Abdominal pain on a 4-point scale ranging from 0 (none) to 3 (severe)
- General well-being on a 5-point scale ranging from 0 (generally well) to 4 (terrible)

FACIT-F

The FACIT-F Scale (Version 4) is a 13-item instrument that measures fatigue and its impact on daily functions over a recall period of 7 days. Five of the items assess the experience of fatigue, and 8 items assess the impact of fatigue. Items are scored on a 5-point verbal rating scale, yielding a score ranging from 0 to 52, with lower scores indicating greater fatigue (see Appendix 7.6).

4.2.8.2 Derivations

Baseline and Change from Baseline are defined in Section 3.3.1.2.

Average stool frequency is calculated as the average number of entries in the BMD in the 7 days prior to the Visit. For visits where ileocolonoscopy is performed, averages will be generated within the 7 days prior to bowel prep. The endpoint is Change from Baseline.

For each stool type Stool, Blood, and Mucus, the percentage for a day is calculated by dividing the number of responses (Stool=[1], Blood=[2], Mucus=[3]) by the total number of stools for the day, multiplied by 100.

Average daily loose stool frequency (LSF) score is calculated as the average daily count of stools of Type 6 or 7 on BSFS during each week. That is, add the total number of stools of Type 6 or Type 7 for each day, then take the mean across all completed days in the 7 days prior to the Visit. For visits where ileocolonoscopy is performed, averages will be generated within the 7 days prior to bowel prep. The endpoint is Change from Baseline.

Percentage of stools with consistency 6 or 7 for each day will be number of responses to question E06 (Table 7-4) that equal 6 or 7, divided by the number of stools that day multiplied by 100. The average will be taken from the 7 days prior to the Visit. For visits where ileocolonoscopy is performed, averages will be generated within the 7 days prior to bowel prep. The endpoint is Change from Baseline.

Urgency for the day is calculated as the number of stools in which the answer to "Did you experience urgency (a sudden, almost irresistible need) to pass this bowel movement?" = [1]. The percentage for a day is calculated by dividing the number of urgent stools by the total number of stools for the day, multiplied by 100. The average will be taken from the 7 days

prior to the Visit. For visits where ileocolonoscopy is performed, averages will be generated within the 7 days prior to bowel prep. The endpoint is Change from Baseline.

Evening diary NRS responses (abdominal pain, fatigue, tiredness, weakness, lack of energy, and joint pain) range from 0 to 10. Calculate the average score for each Visit over the last 7 days prior to the respective Visits. For visits where ileocolonoscopy is performed, averages will be generated within the 7 days prior to bowel prep. The endpoint is Change from Baseline.

Evening diary CDAI responses (LSF, abdominal pain, general well-being) are calculated for each Week by summing for the scores for the 7 day window and dividing by the number of days the diary was completed. For visits where ileocolonoscopy is performed, averages will be generated within the 7 days prior to bowel prep. The endpoint is Change from Baseline.

FACIT-F is completed on site, recalling the past 7 days. It is scored by reversing all items except AN5 and AN7, adding the available item scores, multiplying by 13, and divided by the number of items answered. (FACIT-F.org). The endpoint is Change from Baseline.

4.2.8.3 Handling of Dropouts and Missing Data

BMD, and Evening Diary items

If there are not 4 (consecutive or non-consecutive) days of data available for the week preceding a Visit (or prior to bowel prep for visits where ileocolonoscopy is performed), the window will be extended as described for CDAI PRO items in Section 4.2.1.3. If 4 (consecutive or non-consecutive) days of data are still not available, the value will be set to missing and their data will not be included in the summary table.

FACIT-F Fatigue

If there are missing items on the FACIT-F, and there are at least 50% of items completed, the subscore can be imputed (Webster, Cella, and Yost 2003). Multiply the sum of the responses by the number of items (13), then divide by the number of items answered:

score = [(sum of item scores) * (13)] / (number of items answered)

If more than 50% of the items are missing, then then FACIT-F score will be set to missing. Missing data will be handled in the mixed model.

4.2.8.4 Primary Analysis of Exploratory Endpoints

The average score over 7 days prior to each Visit (or prior to bowel prep for visits where ileocolonoscopy is performed), will be calculated for each endpoint and the Change from Baseline calculated for each Visit for each score. Using the FAS, each measure will be analyzed separately using a mixed model as described in Section 4.2.7.4.

4.2.8.4.1 Presentation

The Change from Baseline at each Visit will be presented using descriptive statistics for each score. Mixed model results for all Visits will be presented as in Section 4.2.7.4.1.

4.2.8.5 Subgroup Analysis

The same analysis as described in Section 4.2.7.5 will be performed for Baseline serum IL-22 (High/BM+, Low/BM-), using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5).

4.2.9 Exploratory Endpoints: Health Related Quality of Life (HRQoL) (IBDQ, SF-36, EQ-5D-5L)

4.2.9.1 Definitions

Inflammatory Bowel Disease Questionnaire (IBDQ)

The IBDQ is presented in Appendix I 3.4 of the CSP, and measures HRQoL in patients with inflammatory bowel disease (IBD). It will be assessed during the site visit at Baseline and Weeks 12, 24, 40, 52, and ET or E/D. It covers how the participant has been feeling in the past two weeks (Guyatt et al, 1989). Four dimensions are covered: bowel symptoms (10 items), systemic symptoms (5 items), emotional function (12 items), and social function (5 items). Items are scored on a 7-point Likert scale (from 1 [worse] to 7 [better]). The total score is obtained by summing the score from each item, yielding a global score in the range 32 to 224 (with higher scores indicating better quality of life).

EQ-5D-5L

The EQ-5D-5L is presented in Appendix I 3.5 of the CSP, and will be assessed at Baseline and during site visits at Weeks 12, 24, 52, and ET or E/D. The EQ-5D-5L includes a Visual Analog Scale (VAS) that allows respondents to rate their own current health on a 101-point scale ranging from “best imaginable” (100) to “worst imaginable” (0) health.

SF-36v2

The SF-36v2 (SF36) is presented in Appendix I 3.6 of the CSP, and will be assessed at Baseline and during site visits at Weeks 12, 24, 52, and ET or E/D. It is a standardized instrument used to measure self-reports of health status and functional well-being. It consists of 8 domains: physical functioning (10 questions, #3a to 3j), role physical (4 questions, #4a to 4b), bodily pain (2 questions, #7 and 8), general health (5 questions, #1,11a to 11d), vitality (4 questions, #9a, 9e, 9g, 9i), social functioning (2 questions, #6, #10), role emotional (3 questions, #5a to 5c), and mental health (5 questions, #9b to 9d, 9f, 9h). The subscores from the 8 dimensions can be directly calculated. Two summary indices can be scored: the Physical

Component Score (PCSc) and the Mental Component Score (MCSc). The lower the score, the more disability (Brazier, Roberts, and Deverill, 2002).

4.2.9.2 Derivations

IBDQ

The total score is obtained by summing each item, and will range from 32 to 224. Change from Baseline is calculated as described in Section 3.3.1.2. A positive Change from Baseline value indicates increased quality of life compared to Baseline.

The mapping of items to the 4 subscores is found in Table 4-7 along with the maximum score per dimension, calculated by summing the response to each item in the subscore.

Table 4-7 IBDQ subscores

Subscore	*Item numbers*	*Max points*
Bowel symptoms	1,5,9,13,17,20,22,24,46,29	70
Systemic symptoms	2,6,10,14,18	35
Emotional function	3,7,11,15,19,21,13,15,17,30,31,32	84
Social function	4,8,12,16,28	35

Each of the 4 IBDQ domain scores is calculated by summing the responses from the items in that domain, as defined in Table 4-7. The average per item scores can be calculated by dividing the each subscore by the number of items in the subscore.

EQ-5D-5L

Change from Baseline is calculated for the 101-point scale. A negative Change from Baseline value indicates a worse health state.

SF-36v2

The SF36 will be scored using the algorithm provided by Optum with the instrument license (Health Outcomes™ Scoring Software 4.5 or higher). Change from Baseline is calculated for the PCSc and MCSc.

4.2.9.3 Handling of Dropouts and Missing Data

IBDQ, EQ-5D-5L, and SF36 will be administered at the site. If a participant discontinues early or misses a Visit, the scores for that Visit will be set to missing and the mixed model will use the non-missing data to estimate the means.

If more than half of the items in a subscore are missing for the IBDQ, the total score will be set to missing as well as that subscore. Otherwise, substitution of the mean of the remaining items in the subscore can be imputed for that subscore (W. Y. Cheung, 2002).

Per the EQ-5D-5L user guide (EQ-5D-5L User Guide, 2019), missing values should be coded as '9', or '999' for the VAS. These should be set to missing prior to analysis.

The SF36 is scored by summing items in the same subscale. If a question is missing, values will be imputed according to the algorithm provided by Optum.

4.2.9.4 Primary Analysis of Exploratory Endpoints

Using the FAS, Change from Baseline for each endpoint (IBDQ, IBDQ subscores, EQ-5D-5L VAS, SF36 PCSc, and SF36 MCSc) will be analyzed via a MMRM as described in Section 4.2.7.4.

The 5 domains of the EQ-5D-5L will not be analyzed.

4.2.9.4.1 Presentation

Absolute values and change from Baseline in IBDQ, EQ-5D-5L VAS, SF36 PCSc, and SF36 MCSc will be presented using descriptive statistics by visit and treatment for the FAS.

For each mixed model, results will be presented as described in Section 4.2.7.4.1, as well as descriptive statistics for each score.

4.2.9.5 Subgroup Analysis

The same analysis as described in Section 4.2.7.5 will be performed for Baseline serum IL-22 (High/BM+, Low/BM-), using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5).

4.3 Pharmacodynamic Endpoints

Individual PD parameters will be listed with Change from Baseline and percentage Change from Baseline by treatment group.

4.3.1 Definitions

Pharmacodynamic (PD) variables collected for this study include IL-22, serum CRP, and FCP.

4.3.2 Handling of Dropouts and Missing Data

There will be no imputation of missing data. If any IL-22, CRP, or FCP are missing for a Visit, then Change from Baseline and percentage Change from Baseline will be set to missing and not be used in the analysis.

4.3.3 Primary Analysis of Exploratory Endpoints

Analysis of IL-22 is described in Section 4.2.5.

Using the FAS, Change from Baseline for each PD parameter (IL-22, CRP, FCP) will be analyzed via a MMRM as described in Section 4.2.7.4.

4.3.3.1 Presentation

Individual PD parameters will be listed with Change from Baseline and percentage Change from Baseline by treatment group. Absolute values, Change from Baseline, and percentage Change from Baseline will be summarized by treatment group and BM status using descriptive statistics.

For each mixed model, results will be presented as described in Section 4.2.7.4.1.

Mean absolute values, mean Change from Baseline, and mean percentage Change from Baseline will also be presented graphically by treatment group.

4.3.4 Subgroup Analysis

The analysis described in Section 4.2.7.4 will be performed by Baseline serum IL-22 (High/BM+, Low/BM-) for the three PD parameters, using the cut-off as determined in the interim (primary) analysis for Stage 1 (see Section 4.2.5).

4.4 Pharmacokinetics

Serum samples for determination of brazikumab concentrations and the presence of anti-drug antibodies (ADA) will be collected at Baseline prior to first study intervention administration, at multiple time points during the Double-Blind Treatment Period, and at the end of the follow-up period. All PK analyses will be based on the PK analysis set defined in Section 3.2.

Individual brazikumab serum concentrations for each nominal sampling time will be listed and summary descriptive statistics will be presented and plotted versus time. Individual brazikumab serum concentrations will be graphically illustrated and summarized by ADA status.

For each time point, the geometric mean, geometric standard deviation, coefficient of variation (CV), arithmetic mean, SD, median, minimum, and maximum will be reported as summary statistics.

Predose PK blood samples will be collected at Weeks 0, 4, 8, 12, 16, 24, and 40. Post-dose PK blood samples will be collected at Week 0, 1 hour and 2-4 hours post-dose, and at Week 8, 1 hour post-dose. Data will be assigned to analysis windows as defined in Section 3.3.2.

Linear and log-linear geometric mean (± SE) concentration profile plots against time will be produced for each treatment arm, with nominal sampling time on the x-axis.

Further, model-based evaluation of PK will be detailed in a separate PK analysis plan and reported outside the CSR.

4.5 Immunogenicity

Serum samples to measure the presence of ADAs against brazikumab (binding antibodies and neutralizing antibodies) will be collected at Weeks 0, 12, 24, 40, and 52. ADA results from each sample will be reported as either positive or negative. If the sample is positive, the ADA titre will be reported as well. In addition, the presence of nAb will be tested for all ADA-positive samples.

4.5.1 Definitions

ADA positive/negative status at the sample level:

- A sample is considered to be ADA positive if the titre is greater than or equal to the minimum required dilution (MRD) of the assay.
- A sample is considered to be ADA negative if the titre is <MRD of the assay.

ADA positive/negative status at the subject level:

- A subject is considered ADA positive if a collected sample is tested positive at any time during the study, including baseline and/or post-baseline (see definition for ADA prevalence below).
- A subject is considered ADA negative if collected samples are tested negative at all timepoints, including baseline and post-baseline.

Treatment-related ADA development at subject level:

- Treatment-emergent ADA positive (TE-ADA+): A positive post-baseline result and either of the following statements holds (Shankar et al, 2014):
  - Baseline is ADA negative and at least one post-baseline assessment is ADA positive. This is called treatment-induced ADA positive.
  - Baseline is ADA positive, and the baseline titre is boosted by greater than the variability of the assay (i.e. >= X-fold increase, commonly 4-fold) at >=1 post-baseline timepoint. This is called treatment-boosted ADA positive.
- Non-treatment-emergent ADA positive (non-TE-ADA+): Subjects who are ADA positive but not fulfilling the conditions for TE-ADA+

4.5.2 Derivations

The number and percentage of ADA-evaluable participants in the following ADA categories in each treatment group and overall will be determined. The number of ADA-evaluable participants in the treatment group will be used as the denominator for percentage calculation.

- Subjects who are ADA positive at any time during the study, including baseline and/or post-baseline (also generally referred to as ADA positive). The proportion of ADA-positive subjects in a population is known as ADA prevalence.

- Subjects who are treatment-emergent ADA positive (see definition in Section 4.5.1), reported overall and separately as treatment-induced and treatment-boosted subjects. The proportion of TE-ADA+ subjects in a population is known as ADA incidence.
- Subjects who are non-treatment-emergent ADA positive (see definition in Section 4.5.1)
- Subjects who are ADA positive at baseline and at at least one post-baseline assessment.
- Subjects who are ADA positive at baseline only.
- Subjects who are persistently ADA positive, which is defined as ADA negative at baseline and having at least 2 post-baseline ADA positive measurements with >= 16 weeks between first and last positive, or an ADA positive result at the last available post baseline assessment.
- Subjects who are transiently ADA positive, defined as ADA negative at baseline and having at least one post-baseline ADA positive measurement and not fulfilling the conditions for persistently positive.
- Subjects who are ADA positive with maximum post-baseline titre > median of maximum post-baseline titres.
- Subjects who are neutralizingg antibodies (nAb) positive at any time during the study, including baseline and/or post-baseline (also generally referred to as nAb positive). The proportion of nAb-positive subjects in a population is known as nAb prevalence.
- Subjects who are nAb negative at baseline (or ADA negative at baseline) and nAb positive at any post-baseline visit. The proportion of such subjects in a population is known as nAb incidence. Note: Subjects who are ADA negative at baseline are included to ensure that all subjects who are nAb positive for the first-time post-baseline satisfy this definition, given that subjects who are ADA negative at baseline would not have a nAb result reported.

4.5.3 Primary Analysis of ADA

A summary of the number and percentage of participants together with descriptive statistics of the maximum titre for each ADA category (Section 4.5.2) in different treatment groups and overall will be presented based on the ADA evaluable analysis set. In addition, descriptive statistics of the titre for ADA positive participants will be presented over time.

Additional immunogenicity outcomes may be described in separate analysis plans and will be reported separately.

4.6 Safety Analyses

Safety data will be summarized descriptively using the Safety analysis set, which includes participants who enrolled prior to and after the inititiation of CSP Amendment 4. Pre-Amendment 4 participants randomized to the Humira® group will be listed only. The other Pre-Amendment 4 participants will be grouped with their corresponding Post-Amendment 4 dose group and included in safety summary tables.

The safety parameters will include AEs, clinical laboratory, vital signs, physical examinations, and ECG parameters. Continuous variables will be summarized by the number of participants and mean, SD, median, minimum, and maximum values. Categorical variables will be summarized by number and percentage of participants. All safety parameters, except AEs, will be summarized for the Induction Period and the Maintenance Period separately.

4.6.1 Exposure

4.6.1.1 Definitions and Derivations

Exposure to the study intervention will be calculated as the number of days from the first study intervention to 28 days after the last study intervention for the study for the Safety analysis set. Any erroneous exposure to brazikumab in the placebo group will be counted as a brazikumab exposure. Any gaps in dosing will be ignored when calculating the total.

Total dose will be calculated as the IV doses given during the Induction Period during Week 0, 4, and 8, plus the SC doses given from Week 12 through Week 48.

4.6.1.2 Presentation

Exposure to the study intervention (in days) and total dose (in mg) during the Double-Blind Treatment Period will be listed and summarized by treatment group for the Safety analysis set.

4.6.2 Adverse Events

Adverse events (AE) will be coded by system organ class and preferred term using the MedDRA, version 24.0 or newer.

4.6.2.1 Definitions and Derivations

Adverse event data will be categorized according to their onset date based on the last dose of investigational product as follows:

- An AE occurring during the Treatment Period is defined as an AE with a date of onset ≥ the first dose of investigational product and ≤ 28 days after the last dose of investigational product

- An AE occurring during the Study Period is defined as an AE that occurs during treatment including follow-up, with a date of onset ≥ the first dose of investigational product and ≤ 18 weeks after the last dose of investigational product

Treatment related

An AE will be considered as treatment related if the investigator considers that there is a reasonable possibility that the event may have been caused by the investigational product.

Serious

An AE will be considered a serious AE (SAE) if it is a AE that additionally meets any SAE criteria (as recorded on the AE form of the eCRF).

Missing Severity Assessment

Adverse events with missing intensity are assumed to be severe. The imputed values for severity assessment will be used for the incidence summary; the values will be shown as missing in the data listings.

Missing Causal Relationship to Study Intervention

There will be no imputation of missing causal relationship for any non-serious AEs. For SAEs, if the causal relationship to the study intervention is missing for an SAE that started on or after the date of the first dose of the study intervention, a causality of yes will be assigned. The imputed values for causal relationship during the Treatment Period will be used for the incidence summary; the values will be shown as missing in the data listings.

Missing Date Information for Adverse Events

See Appendix 7.8 for missing date imputation rules.

4.6.2.2 Presentation

Adverse Events will be reported for the Treatment Period and Study Period (Treatment Period + Safety Follow-up Period). The end of the Treatment Period is defined as the date of last dose of study intervention + 28 days. The end of the Follow-up Period is defined as the date of last dose of study intervention + 18 weeks.

For the Safety analysis set, the following will be summarized by treatment group and overall:

- Overall summary of AE, including
  - Any AE
  - Any SAE
  - Any SAE with outcome death
  - Any AE leading to discontinuation of IP
  - Any possibly related AE

- Any possibly related SAE

- AE by SOC and PT
- AE sorted by decreasing frequency on preferred term level
- AE by maximum reported intensity on preferred term level
- Possibly related AE by SOC and PT
- AE per 100 subject years by SOC and PT
- Adverse events of special interest (AESI)
- SAE with outcome death by SOC and PT
- SAE by SOC and PT
- AE leading to discontinuation of investigational product by SOC and PT

The incidence of common (> 5% of participants in either 720 mg IV, 1440 mg IV, 120 mg SC, or 240 mg SC treatment groups) non-serious AEs will be summarized by system organ class, preferred term, and treatment group.

If more than 1 AE is coded to the same preferred term for the same participant, the participant will be counted only once for that preferred term using the most severe severity for the summaries by severity. Similarly, the participant will be counted using the closest relationship for summaries by causality.

Listings of AEs, deaths, SAEs and AEs leading to discontinuation of the investigational product by participant will be presented, to include key participant information, including Subject identifier, sex, age, race, event term as reported, AE (PT), time from first dose to AE, treatment period, last dose prior to event, time from first dose to <death or discontinuation>, intensity, seriousness, whether treatment was received for the AE, action taken with respect to study intervention, outcome, and AE relationship to study intervention.

AESI will be summarized and include infusion related reactions/injection-site reactions, malignancies, hypersensitivity (including anaphylaxis), and infections.

4.6.3 Clinical Laboratory, Blood Sample

4.6.3.1 Definitions and Derivations

The following are protocol-required laboratory parameters as listed in Table 4-8 through Table 4-10. All laboratory assessments will be done at Screening. Clinical chemistry and hematology will be carried out at Weeks 0, 4, 12, 24, 40, 52, unscheduled visits, E/D or ET, and safety follow-up visits. Hematocrit and CRP will be assessed at every visit up to Week 12, then at Weeks 24, 40, 52, unscheduled visits, E/D or ET, and safety follow-up visits. Handling of unscheduled, E/D or ET visits is described in Section 3.3.3.

Table 4-8 Protocol-Required Laboratory Assessments

Laboratory Assessments	***Parameters***			
Hematology	Platelet count RBC count Hemoglobin Hematocrit	RBC indices: MCV MCH % Reticulocytes		WBC count with differential: Neutrophils Lymphocytes Monocytes Eosinophils Basophils
Clinical Chemistry	BUN C-reactive protein	Potassium	AST	Total and direct bilirubin
	Creatinine eGRF	Sodium	ALT	Total protein
	Glucose (nonfasting)	Calcium	Alkaline phosphatase	Bicarbonate
	Magnesium	Chloride	Albumin	Phosphate
	Uric acid			
Fecal tests	• *Clostridium difficile* • Fecal calprotectin • Fecal lactoferrin			
Routine urinalysis	• Specific gravity • pH, glucose, protein, blood, ketones, bilirubin, urobilinogen, nitrite, leukocyte esterase by dipstick • Microscopic examination (reflexively as needed)			
Other Screening tests	• QuantiFERON-TB Gold In-Tube • FSH (as needed in women of non-childbearing potential only) • Serum hCG pregnancy test (as needed for women of childbearing potential). Urine hCG is to be performed locally at each visit prior to administering study intervention and as outlined in the SoA. • Serology (HIV antibody, hepatitis B surface antigen, hepatitis B core total antibody, hepatitis B core IgM antibody, and hepatitis C virus antibody with reflex testing when required) • All study-required laboratory assessments will be performed by a central laboratory unless indicated (eg, local urine hCG testing)			

ALT = alanine aminotransferase; AST = aspartate aminotransferase; BUN = blood urea nitrogen; eGFR = estimated glomerular filtration rate; FSH = follicle stimulating hormone; hCG = human chorionic gonadotropin; HIV = human immunodeficiency virus; IgM = immunoglobulin M; MCH = mean corpuscular hemoglobin; MCV = mean corpuscular volume; RBC = red blood cell; SoA = Schedule of Activities; TB = tuberculosis; WBC = white blood cell.

Table 4-9 Normal ranges for laboratory values (Central Laboratory Services Manual)

Laboratory Parameter	*LLN–- Male*	*ULN- Male*	*LLN – Female*	*ULN- Female*	*SI Unit*
Hematology					
Hemoglobin	Age 12-59: 127 Age >59: 125	Age 12-59: 181 Age >59:170	Age 12-59: 116 Age >59: 115	Age 12-59: 164 Age >59: 158	g/L
Hematocrit	Age 12-59: 0.39 Age >59: 0.37	Age 12-59: 0.54 Age >59: 0.51	0.34	0.48	ratio
RBC	Age 12-59: 4.5 Age >59: 4.0	Age 12-59: 6.4 Age >59: 5.8	Age 12-59: 4.1 Age >59: 3.9	Age 12-59: 5.6 Age >59: 5.5	TI/L
Eosinophils	0.00	0.57	0.00	0.57	GI/L
Neutrophils	40.5	75.0	40.5	75.0	%
Basophils	0.00	0.20	0.00	0.20	GI/L
Monocytes	0.12	0.92	0.12	0.92	GI/L
Lymphocytes	Age 18-59: 0.91 Age >59: 0.80	Age 18-59: 4.28 Age >59:3.00	Age 18-59: 0.91 Age >59: 0.80	Age 18-59: 4.28 Age >59: 3.00	GI/L
Absolute neutrophil count	1.96	7.23	1.96	7.23	GI/L
Platelet count	Age 12-60: 140 Age > 60: 130	Age 12-60: 400 Age > 60: 394	Age 12-60: 140 Age > 60: 130	Age 12-60: 400 Age > 60: 394	10^9/L
Clinical chemistry					
Albumin	Age 18-69: 33 Age >69: 33	Age 18-69: 49 Age >69: 46	Age 18-69: 33 Age >69: 33	Age 18-69: 49 Age >69: 46	g/L
Alkaline phosphatase	Age 18-19: 55 Age >19: 40	Age 18-19: 149 Age > 19: 129	Age 18-19: 45 Age > 19: 35	Age 18-19: 87 Age > 19: 104	U/L
Alanine aminotransferase (ALT)	Age 18-69: 6 Age > 69: 6	Age 18-69: 43 Age > 69: 35	Age 18-69: 6 Age > 69: 6	Age 18-69: 34 Age > 69: 32	U/L
Aspartate aminotransferase (AST)	11	36	9	34	U/L
Gamma-glutamyl transferase (GGT)	Age 18-59: 10 Age >59: 10	Age 18-59: 61 Age >59: 50	Age 18-59: 4 Age >59: 5	Age 18-59: 49 Age >59: 50	U/L
Blood urea nitrogen or Urea	1.4	Age 18-70: 8.6 Age >70: 10.4	1.4	Age 18-70: 8.6 Age >70: 10.4	mmol/dL
Calcium	2.07	2.64	2.07	2.64	mmol/dL
Chloride	94	112	94	112	mmol/L
Creatinine	40	Age 18-50: 110 Age 50-70: 119 Age >70: 137	31	Age 18-70: 101 Age >70: 110	µmol/L

Laboratory Parameter	*LLN–- Male*	*ULN- Male*	*LLN – Female*	*ULN- Female*	*SI Unit*
Glucose, nonfasting	3.9	5.6	3.9	5.6	mmol/L
Magnesium	Age 18-20: 0.70 Age 20-60: 0.66 Age >60: 0.66	Age 18-20: 0.90 Age 20-60: 1.07 Age >60: 0.98	Age 18-20: 0.70 Age 20-60: 0.66 Age >60: 0.66	Age 18-20: 0.90 Age 20-60: 1.07 Age >60: 0.98	mmol/L
Potassium	3.5	5.2	3.5	5.2	mmol/L
Sodium	Age 18-59: 132 Age >59: 135	Age 18-59: 147 Age >59: 145	Age 18-59: 132 Age >59: 135	Age 18-59: 147 Age >59: 145	mmol/L
Total bilirubin	3	21	3	21	µmol/L
Total protein	Age 18-59: 61 Age >59: 60	Age 18-59: 84 Age >59: 80	Age 18-59: 61 Age >59: 60	Age 18-59: 61 Age >59: 60	g/L
Uric acid or Urate	Age 18-50: 125 Age > 50: 149	Age 18-50: 488 Age >50: 494	Age 18-50: 125 Age >50: 149	Age 18-50: 428 Age >50: 446	mg/dL

Blood urea nitrogen or Urea are the same parameters, and Uric acid or Urate are the same parameters.
LLN = lower limit of normal; ULN = upper limit of normal laboratory reference range.

Table 4-10 Criteria for Potentially Clinically Significant Laboratory Values

Laboratory Parameter	*SI Unit*	*Conversion Factor*[a]	*Conventional Unit*	*PCS Criterion*[b] *Low Value*	*PCS Criterion*[b] *High Value*
Hematology					
Hemoglobin	g/L	0.1	g/dL	< 0.9 × LLN	—
Hematocrit	Volume fraction	100	%	< 0.9 × LLN	—
Eosinophils	%	1	%	—	> 10
Neutrophils	%	1	%	< 30	> 90
Basophils	%	1	%	—	> 6
Monocytes	%	1	%	—	> 20
Lymphocytes	%	1	%	< 10	> 60
Absolute neutrophil count	× 10^9/L	1	1000/µL	< 1.0	—
Platelet count	× 10^9/L	1	1000/µL	≤ 75	≥ 700
White blood cell count with differential	× 10^9/L	1	1000/µL	≤ 2.5	≥ 15
Chemistry					
Albumin	g/L	0.1	g/dL	< 0.9 × LLN	> 1.1 × ULN
Alkaline phosphatase	U/L	1	U/L	—	≥ 3 × ULN

Laboratory Parameter	*SI Unit*	*Conversion Factor*[a]	*Conventional Unit*	*PCS Criterion*[b] *Low Value*	*PCS Criterion*[b] *High Value*
Alanine aminotransferase (ALT)	U/L	1	U/L	—	≥ 3 × ULN
Aspartate aminotransferase (AST)	U/L	1	U/L	—	≥ 3 × ULN
Gamma-glutamyl transferase (GGT)	U/L	1	U/L	—	≥ 3 × ULN
Blood urea nitrogen or Urea	mmol/L	2.8011	mg/dL	—	> 1.2 × ULN
Calcium	mmol/L	4.008	mg/dL	< 0.9 × LLN	> 1.1 × ULN
Chloride	mmol/L	1	mg/dL	< 0.9 × LLN	> 1.1 × ULN
Creatinine	µmol/L	0.0113	mg/dL	—	> 1.3 × ULN
Glucose, nonfasting	mmol/L	18.018	mg/dL	< 0.8 × LLN	> 1.2 × ULN
Magnesium	mmol/L	2	mEq/L	< 0.9 × LLN	> 1.1 × ULN
Potassium	mmol/L	1	mEq/L	< 0.9 × LLN	> 1.1 × ULN
Sodium	mmol/L	1	mEq/L	< 0.9 × LLN	> 1.1 × ULN
Total bilirubin	µmol/L	0.0585	mg/dL	—	> 1.5 × ULN
Total protein	g/L	0.1	g/dL	< 0.9 × LLN	> 1.1 × ULN
Uric acid or Urate	µmol/L	0.0168	mg/dL	—	> 1.1 × ULN

[a] Conversion factor from SI units to conventional (traditional) units.
[b] Criteria refer to SI units.
Blood urea nitrogen or Urea are the same parameters, and Uric acid or Urate are the same parameters.
LLN = lower limit of normal; PCS = potentially clinically significant; SI = *Le Système International d'Unités* (International System of Units); ULN = upper limit of normal laboratory reference range.

Clinical laboratory test values will be considered potentially clinically significant (PCS) if they meet either the lower-limit or higher-limit PCS criteria listed in Table 4-10 (using the same periods as defined for AEs in Section 4.6.2.1).

4.6.3.2 Presentations

Observed, and Change from Baseline for Hematology and clinical chemistry parameters will be listed and summarized in SI units by treatment group for the Safety analysis set.

If the reported value of a clinical laboratory parameter cannot be used in a statistical summary table because, for example, a character string is reported for a parameter of the numeric type, a coded value must be appropriately determined for use in the statistical analyses. The actual values, however, as reported in the database will be presented in the data listings.

The number and percentage of participants who have PCS post baseline clinical laboratory values will be tabulated by treatment group for the Double-Blind Treatment Period. The

percentages will be calculated relative to the number of participants with available baseline values and at least 1 post baseline assessment for the Double-Blind Treatment Period. The numerator will be the total number of participants with available baseline values and at least 1 PCS post baseline value for the Double-Blind Treatment Period. A supportive tabular display of participants with PCS post baseline values will be provided, including the participant identified, baseline and all post baseline (including non-PCS) values.

Chemistry lab results will be similarly listed and summarized with descriptive statistics.

4.6.4 Clinical Laboratory, Urinalysis

4.6.4.1 Definitions and Derivations

Urinalysis will be carried out at Screening and at Weeks 0, 4, 12, 24, 40, 52, unscheduled visits, E/D or ET, and safety follow-up visits. Handling of unscheduled, E/D or ET visits is described in Section 3.3.3.

Routine urinalysis includes specific gravity, pH, glucose, protein, ketones, bilirubin, urobilinogen, nitrite, and leukocyte esterase, all of which are considered a high value if 2+. In addition, pH normal range is ≥ 5 to ≤ 8. Microscopic examination is performed as needed.

4.6.4.2 Presentation

Listing of urinalysis results, for the Safety analysis set will be provided. A shift table from Baseline to the end of the Treatment Period will be presented by treatment group. Treatment emergent urinalysis abnormalities will be summarized by treatment group.

4.6.5 Other Laboratory Evaluations

4.6.5.1 Definitions and Derivations

Other laboratory testing is presented in Table 4-8.

4.6.5.2 Presentations

Positive pregnancy and other screening tests will be listed.

4.6.6 Physical Examinations

4.6.6.1 Definitions and Derivations

A complete physical examination will include assessments of the following:

- general appearance
- respiratory

- cardiovascular
- abdomen
- skin
- head and neck (including ears, eyes, nose, and throat)
- lymph nodes
- thyroid
- muscular-skeletal (including spine and extremities)
- neurological systems
- fistula exam (as applicable)

Physical examinations will be performed at Screening, Weeks 0, 2, 4, 8, 12, 24, 40, 52, unscheduled visits, E/D or ET, and safety follow-up visits.

Abdominal mass and extraintestinal manifestations (EIM) will be assessed during the physical examination and will be used to calculate CDAI.

Any new findings or aggravated existing abnormalities, judged as clinically significant by the investigator, will be reported as an AE.

4.6.7 Vital Signs

4.6.7.1 Definitions and Derivations

Vital signs will be collected at every Visit and include:

- Pulse rate
- Respiratory rate
- Temperature
- Blood pressure
- Height (Screening Visit 1 only) recorded in centimeters
- Weight, recorded in kilograms

Vital sign values will be considered PCS if they meet both the observed value criterion and the change from baseline value criterion, if both criteria are available, or meet either the observed value criterion or the change from baseline value criterion detailed in Table 4-11.

Table 4-11 Criteria for Potentially Clinically Significant Vital Signs

Parameter	***Flag***	***Criteria[a]***	
		Observed Value	*Change from Baseline*
Sitting systolic blood pressure, mm Hg	High	≥ 140	Increase of ≥ 20
	Low	<90	Decrease of ≥ 20
Sitting diastolic blood pressure, mm Hg	High	≥ 90	Increase of ≥ 15
	Low	<60	Decrease of ≥ 15
Sitting pulse rate, bpm	High	≥ 100	Increase of ≥ 15
	Low	≤ 60	Decrease of ≥ 15
Weight, kg	High	—	Increase of ≥ 7%
	Low	—	Decrease of ≥ 7%

[a] A post baseline value is considered potentially clinically significant if it meets both the observed-value and the change-from-baseline criteria.
bpm = beats per minute; hG = Mercury.

4.6.7.2 Presentation

A listing and summary of descriptive statistics for vital signs (systolic and diastolic blood pressures, pulse rate, body temperature, respiratory rate, and weight) values at Baseline, post baseline, and change from baseline values at each post baseline timepoint will be presented by treatment group for the Safety analysis set.

The number and percentage of participants who have PCS post baseline vital sign values will be tabulated by treatment for each assessment for the Double-Blind Treatment Period. The percentages will be calculated relative to the number of participants who have available baseline values and at least 1 post baseline assessment for the double-blind treatment period. The numerator will be the total number of participants with at least 1 PCS post baseline value for the Double-Blind Treatment Period. Key participant information for participants with PCS post baseline values will be provided.

4.6.8 Electrocardiogram

4.6.8.1 Definitions and Derivations

Single 12-lead ECGs will be obtained at Screening, Weeks 0, 12, 24, 40, 52, E/D or ET, and safety follow-up visits using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QTc intervals. The upper limit for each measure is shown in Table 4-12.

Table 4-12 Criteria for Potentially Clinically Significant Electrocardiograms

Parameter	*Unit*	*Low Value*	*Low Decrease*	*High Value*	*High Increase*
ECG heart rate	bpm	≤ 60	≥ 15	≥ 100	≥15
RR interval	msec	< 500	NA	> 1500	NA
PR interval	msec	NA	NA	≥ 240	NA
QRS duration	msec	≤ 60	NA	≥ 140	NA
QT	msec	≤ 300	NA	≥ 450	> 60
QTcF	msec	≤ 300	NA	≥ 480	> 60
QTcB	msec	≤ 300	NA	≥ 480	> 60

QTc = QT interval corrected for heart rate.
QTcB = QT interval corrected for heart rate using the Bazett formula.
QTcF = QT interval corrected for heart rate using the Fridericia formula.

The outcome of the ECG evaluation is to be recorded as normal or abnormal in the eCRF, with any abnormalities recorded as either not clinically significant, or clinically significant.

4.6.8.2 Presentation

A listing and summary of descriptive statistics for ECG parameters (i.e., heart rate, PR interval, QRS interval, QT interval, and QTc interval) at Baseline, post baseline, and Change from Baseline values at each post baseline timepoint will be presented by treatment group for the Safety analysis set.

ECG parameter values are considered PCS if ECG values meet the value listed in Table 4-12. The number and percentage of participants with PCS post baseline values will be tabulated by treatment for the Double-Blind Treatment Period. The percentages will be calculated relative to the number of participants with an available non-PCS baseline value and at least one post baseline assessment. The numerator will be the total number of participants with an available non-PCS baseline value and at least one PCS post baseline ECG value.

A supportive listing of participants with PCS post baseline values will be provided and will include the participant number and the baseline and post baseline values.

4.6.9 Other Safety Assessments

4.6.9.1 Definitions and Derivations

Potential Hy's Law criteria is defined as a post baseline elevation of ALT or AST ≥ 3x ULN, and TBL ≥ 2x ULN, where at least one time Total Bilirubin ≥ 2x ULN occurred after the first occurrence of AST ≥ 3x ULN or ALT ≥ 3x ULN after start of treatment with investigational product.

Hy's Law is defined as ALT or AST ≥ 3x ULN together with total bilirubin ≥ 2x ULN, where for no other reason, other than study intervention, can be found to explain the combination

of increases (e.g., elevated alkaline phosphatase indicating cholestasis, viral hepatitis, or another drug).

More information can be found in the CSP, Appendix E.

4.6.9.2 Presentations

Participants with elevated liver tests based on measured laboratory values will be summarized for the Safety analysis set. Participants who meet the potential Hy's Law criteria from the first dose of the study intervention to the end of the Induction Period, and in the Maintenance Period to within 18 weeks after the last dose of the study intervention will be listed and summarized by treatment group for the Safety analysis set.

5 INTERIM/PRIMARY ANALYSIS

Details on how trial integrity will be maintained is documented in a separate Study Integrity Plan.

Interim Analysis (IA), as per ICH E9, is an analysis prior to the time point of the formal completion of the trial. The primary analyses in study D5271C00001 will be performed when all participants have completed their Week 12 visit or discontinued early. Accordingly, the pre-specified primary analysis fulfils the ICH E9 definition of an interim analysis. These analyses are described in Section 4.2.

There will be 2 clinical database locks in the study, one at the time of the primary analyses, and one after the last participant has completed the second safety follow-up visit or discontinued early.

All data collected, including Week 52 will be analyzed at the time of the interim (primary) analysis. All available data from participants who have completed Week 52, or would have completed Week 52 if they hadn't discontinued, will be analyzed.

All analyses that will be performed beyond Week 12 at the time of interim (primary) analyses are exploratory and should be interpreted with caution.

6 REFERENCES

Best WR, Becktel JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology 1976; 70(3):439-44.

Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ. 2002 Mar;21(2):271-92. doi: 10.1016/s0167-6296(01)00130-8. PMID: 11939242.

Buxton MJ, Lacey LA, Feagan BG, Niecko T, Miller DW, Townsend RJ. Mapping from disease-specific measures to utility: an analysis of the relationships between the Inflammatory Bowel Disease

Questionnaire and Croh''s Disease Activity Index in Croh''s disease and measures of utility. Value Health 2007; 10:214-20.

Byrum B, Breedon P, Tulkki-Wilke, Platko JV. Meaningful change: Defining the interpretability of changes in endpoints derived from interactive and mHealth technologies in healthcare and clinical research. J of Rehab & Assistive Tech Eng. 7:205566831989277.

Daperno M, ''Haens G, Van Assche G, Baert F, Bulois P, Maunoury V, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. Gastrointest Endosc. 2004;60:505-512.

EuroQol Research Foundation. EQ-5D-5L User Guide, 2019. Available from: https://euroqol.org/publications/user-guides.

Guyatt G, Mitchell A, Irvine EJ, et al. A New Measure of Health Status for Clinical Trials in Inflammatory Bowel Disease. Gastroenterology 1989;96:804-10.

Irvine EJ. Development and subsequent refinement of the inflammatory bowel disease questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel disease. J Pediatr Gastroenterol Nutr. 1999 Apr;28(4):S23-7.

Loftus EV, Feagan BG, Colombel JF, Rubin DT, Wu EQ, Pollack PF, Chao J, Mulani P. Effects of adalimumab maintenance therapy on health-related quality of life of patients with Crohn's disease: patient-reported outcomes of the CHARM trial. Am J Gastroenterol 2008; 103:3132-3141.

Rubin GP, Hungin APS, Chinn DJ, Dwarakanath D. Quality of life in patients with established inflammatory bowel disease: a UK general practice survey. Aliment Pharmacol Ther 2004; 19:529-535.

Sandborn, WJ, Vermeire S, Sands, BE, et al. Defining Remission Using Patient-reported Outcome Components of the Croh''s Disease Activity Index: Post-hoc Analyses of Vedolizumab Data from the GEMINI 2 Study, American Journal of Gastroenterology: October 2015–- Volume 110–- Issue–- p S792-S793

Sands BE, Ooi CJ. A survey of methodological variation in the Crohn's disease activity index. Inflamm Bowel Dis. 2005; 11:133-38.

Spiegel B, Bolus R, Harris LA, et al. Measuring irritable bowel syndrome patient-reported outcomes with an abdominal pain numeric rating scale. *Aliment Pharmacol Ther*. 2009;30(11-12):1159-1170. doi:10.1111/j.1365-2036.2009.04144.x.

Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes*. 2003;1:79. Published 2003 Dec 16. doi:10.1186/1477-7525-1-79.

7 APPENDIX

7.1 Crohn's Disease Activity Index

The CDAI is a composite index with weighted domains that quantifies the global disease severity in a single numerical score. The CDAI measures the severity of active disease using symptom scores that are monitored over the previous week and includes participant-reported symptoms, physician-assessed signs, and laboratory markers (Best 1976, Sands 2005). The CDAI score is calculated by summing weighted scores for subjective items (number of liquid or very soft stools [LSF], abdominal pain [AP] and general well-being) recorded by a diary during a 1-week period, and objective items (associated symptoms, taking antidiarrheal such as loperamide/opiates, abdominal mass, hematocrit, daily morning temperature, and body weight).

The components of the CDAI score are shown in Table 7-1:

Table 7-1 Components of the CDAI score

Physical examination	Abdominal mass EIM (peripheral/axial arthropathy, iritis/uveitis, erythema nodosum, pyoderma gangrenosum, aphthous stomatitis/aphthous ulcers, anal fissures, anal abscess, non-anal abscess, fistula, other)
Antidiarrheal medications	Antidiarrheal medications
Laboratory assessment	Hematocrit
Vital signs	Weight
eDiary	Temperature
Patient-reported components	LSF AP General well-being

7.2 Participant-completed PRO e-Diary: LSF and AP

Patient Reported Outcomes Questionnaire was removed due to copyrights.

Patient Reported Outcomes Questionnaire was removed due to copyrights.

7.3 Simple Endoscopic Score for Croh''s Disease (SES-CD)

Patient Reported Outcomes Questionnaire: SES-CD was removed due to copyrights.

Patient Reported Outcomes Questionnaire: SES-CD was removed due to copyrights.

7.4 Bowel Movement Diary

Patient Reported Outcomes Questionnaire: Bowel Movement Diary was removed due to copyrights.

Patient Reported Outcomes Questionnaire: Bowel Movement Diary was removed due to copyrights.

7.5 Evening eDiary

Patient Reported Outcomes Questionnaire: Evening eDiary was removed due to copyrights.

Patient Reported Outcomes Questionnaire: Evening eDiary was removed due to copyrights.

7.6 The FACIT-F Scale (Version 4)

Patient Reported Outcomes Questionnaire: FACIT-F Scale was removed due to copyrights.

7.7 Study Conduct Mitigation During Study Disruptions Due to Cases of Civil Crisis, Natural Disaster, or Public Health Crisis

The guidance given below supersedes instructions provided elsewhere in this document and should be implemented only during cases of civil crisis, natural disaster, or public health crisis (eg, during quarantines and resulting site closures, regional travel restrictions, and considerations if site personnel or study participants become infected with SARS CoV-2 or similar pandemic infection), which would prevent the conduct of study-related activities at study sites, thereby compromising the study site staff or the participant's ability to conduct the study. The investigator or designee should contact the study sponsor to discuss whether the mitigation plans below should be implemented.

To ensure continuity of the clinical study during a civil crisis, natural disaster, or public health crisis, changes may be implemented to ensure the safety of study participants, maintain compliance with Good Clinical Practice (GCP), and minimize risks to study integrity. Where allowable by local health authorities, ethics committees, healthcare provider guidelines (eg, hospital policies) or local government, these changes may include the following options:

- Obtaining consent/reconsent for the mitigation procedures (note, in the case of verbal consent/reconsent, the ICF should be signed according to local regulations).
- Rescreening: Additional rescreening for screen failure and to confirm eligibility to participate in the clinical study can be performed in previously screened participants. The investigator should confirm this with the sponsor Study Physician/designee.
- Home or Remote visit: Performed by a site qualified Health Care Professional (HCP) or HCP provided by a Third Party Vendor (TPV).
- Telemedicine visit: Remote contact with the participants using telecommunications technology including phone calls, virtual or video visits, and mobile health devices.
- At-home study intervention administration: Performed by a site qualified HCP or HCP provided by a TPV, if possible. Additional information related to the visit can be obtained via telemedicine.

7.8 Missing start and stop dates

For adverse events and prior or concomitant medications, including rescue medications, incomplete (i.e., partly missing) start dates and/or stop dates will be imputed. When the start date and the stop date are both incomplete for a participant, the start date will be imputed first.

7.8.1 Incomplete Start Date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication start date. If the stop date is complete (or imputed) and the imputed start date is after the stop date, the start date will be imputed using the stop date.

If the start date is missing, then it will be set to the date of the first study intervention, unless the end date suggests that the start date could be prior to the date of the first study intervention. In that case, set the start date to January 1 of the same year as the end date.

Missing month and day
The start date will be set to January 1 of the year of occurrence.

If the start year is the same as the year of the first study intervention, then the start date will be set to the date of first study intervention.

Missing month only
If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure.

Missing day only
The missing start day will be set to the first day of the month of occurrence, if the start month and year is after the month and year of the first study intervention.

If the month and year of the incomplete start date are the same as the month and year of the first dose of the study intervention, the day of the first dose of the study intervention will be assigned to the missing day.

If the month and year of the incomplete start date is before the month and year of the first study intervention, the date of informed consent will be assigned to the incomplete date.

7.8.2 Incomplete Stop Date

Because participants will receive their double-blind study intervention doses under the direct supervision of study center personal, the date of dose of study intervention will always be known. The following rules will be applied to impute missing date fields for an incomplete AE, prior, or concomitant medication stop date. If the imputed stop date is before the start date (imputed or non-imputed start date), the imputed stop date will be equal to the start date.

If the end date is missing and the event is not recorded as ongoing, then set the end date as the date of the first study intervention, if the start date is prior to the date of the first study intervention. If the start date is on or after the date of first study intervention, set the end date to the date of last visit.

Missing month and day

If the year of occurrence is the same as the last study intervention year, then the incomplete date will be set to the date of last study intervention. If the year of the incomplete stop date is the same as the year of the last dose of study treatment, the month and day of the last dose of study treatment will be assigned to the missing fields.

If the year of the incomplete stop date is before the year of the last dose of study treatment, *December 31* of the year of occurrence will be assigned to the missing fields, that is, the stop date will be set to December 31 of the year of occurrence. If the participant died in the same year, the stop date will be set to the date of death.

Missing month only

If only the month is missing, the day will be treated as missing and both the month and the day will be replaced as described in Section 7.8.1.

Missing day only

If the month and year of the incomplete stop date are after the month and year of the first study intervention, then the missing end day will be set to the last day of the month of the occurrence.

If the participant died, the end date will be set to the death date.

If the month and year of the incomplete stop date are the same as the month and year of the last dose of study intervention, the day of the last dose of study intervention will be assigned to the missing day.

If either the year of the incomplete stop date is before the year of the date of the last dose of study intervention or if both years are the same, but the month of the incomplete stop date is before the month of the date of the last dose of study intervention, the last day of the month will be assigned to the missing day.

If either the year of the incomplete stop date is after the year of the date of the last dose of study intervention or if both years are the same, but the month of the incomplete stop date is after the month of the date of the last dose of study intervention, the first day of the month will be assigned to the missing day.

7.9 Differential effect method

David Svensson and Fredrik Ohrn. A technical note on determining a biomarker split for differentiating treatment effect in Subpopulations. February 2021

1 Introduction

In this document we generically discuss an approach that could be useful in a RCT setting when a subgroup with enhanced treatment effect is desired to be found in a data-driven fashion. There is a vast literature on the general case where baseline data for multiple covarriates is available, but we here restrict ourselves to the special case of a single biomarker. This is not so uncommon in the practice, and the question then is if there is some simpler and transparent approach that doesn't rely on Machine Learning and furthermore has reasonable theoretical properties and is relatively non-parametric (in the sense of not assuming any particular functional form).

We here describe one such idea in mathematical terms. The setting we have in mind is rather general but assume for now the endpoint Y_i to be binary, let T_i denote treatment allocation $\in \{0,1\}$ and x_i the biomarker; here, $i = 1, ..., n$ denotes subjects in the trial. In the standard primary analysis for an overall treatment effect, we assume that some statistical model M is applied to the full population, rendering an estimate $\hat{\theta}$, with associated standard error $\hat{s}$. (Since our biomarker splitting approach is generic and doesn't rely on any particular statistical modelling for the primary analysis, we simply illustrate the key concepts with a standard test for a difference in proportions).

Assume, without loss of generality, that the range of the biomarker x in the trial is $R = (0,1)$ (hence, possibly scaled), and assume furthermore that any possible value $h \in R$ is considered as a candidate split. This would render two candidate subgroups $S_h^a = \{x : x < h\}$ and $S_h^b = \{x : x \geq h\}$. In each of these sub-populations, the model M can be applied, rendering an estimate and the corresponding standard error. Denote these with $\hat{\theta}_a^{(h)}$, $\hat{\theta}_b^{(h)}$, $s(\hat{h})_a$ and $s(\hat{h})_b$. To assess how well the split differentiate in terms of treatment effect hetereogeneity,

1

we suggest the simple metric

$$g(h) = \frac{\hat{\theta}_b^{(h)} - \hat{\theta}_a^{(h)}}{\hat{s}_\Delta}.$$

where $\hat{s}_\Delta^2$ is the variance of the nominator; since the data in $\{x : x < h\}$ and $\{x : x \geq h\}$ is mutually exclusive, independence holds and

$$\hat{s}_\Delta^2 = s(\hat{h})_a^2 + s(\hat{h})_b^2$$

The split h that maximizes the differential effects is now found by

$$argmax_h|\{g(h)\}.|$$

If instead a directional analysis is desired (under an assumption of monotonicity of effect with x), the absolute value can be removed and h is chosen accounting for the underlying directional assumption.

This approach maximizes the differential effects, while penalizing splits where one of the subgroups is small in which case the term $\hat{s}_\Delta^2$ would tend to increase, hence avoiding results with a large point estimate but a very large standard error.

So, this idea simply uses the difference between the effect sizes, but normalizes by the standard error of that difference. (Note that the variance of the difference is the sum of the variances of the two subgroup estimates, since the data is mutually exclusive; i.e., due to independence)

In practice, if the candidate split is too close to the edges, one of the subgroups will have very little data, and possibly only data from one arm is present. To protect numerically against such artefacts, we suggest restricting the analysis to a subset $[a, b]$, $0 < a < b < 1$, where a and b are such that for $h \in [a, b]$, there is sufficient data in both arms to fit the model M for any subgroups resulting from the cut h (to avoid numerical instabilities).

Moreover, in the bulk of the data range, where estimates and standard errors are defined, $g(h)$ will decrease if one of the estimates $\hat{\theta}_a^{(h)}$ and $\hat{\theta}_b^{(h)}$ is uncertain due one of the subgroups being small.

The advantages with this simple non-parametric approach is that it is straightforward to implement, it avoids splits in the extreme ends, and relies on a well known principle in the literature on causal effects and ITE estimation, namely looking for partitions of the data that maximizes the differential effect. This principle underlies many modern ML methods in this area, such as SIDES [1], GUIDE [2], and QUINT [3]. Similarly, discovering splits that maximizes the differential effects was explored in a spline setting in [4].

As a final remark, it is standard in ITE literature to optimize splits with respect to differential effects rather than to directly optimize the effect in one of the subgroup nodes; the latter approach (optimizing only wrt $S(h)_b$) risk to result in low prevalence subgroups with high uncertainty in the effect estimate (unless some manual arbitrary lower bounds are imposed in the search), whereas the former approach balance out this, see e.g., [1].

References

[1] Lipkovich I., Dmitrienko A., Denne J., Enas G. *Subgroup Identification based on Differential Effect Search (SIDES) a recursive partitioning method for establishing response to treatment in patient subpopulations.* Stat. In. Medicine, 2011.

[2] Loh W., He X., Man M. *A Regression Tree Approach to Identifying Subgroups with Differential Treatment Effects Statistics in Medicine.* January 2014.

[3] Dusseldorp E., Mechelen I. *Qualitative interaction trees: a tool to identify qualitative treatment subgroup interactions.* Statistics in Medicine, 2014, 33(2).

[4] Thomas M., Bornkamp B. *Comparing Approaches to Treatment Effect Estimation for Subgroups in Clinical Trials.* Statistics in Biopharmaceutical Research, 9:2, 2017.

SIGNATURE PAGE

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature

Document Name: d5271c00001-sap-legacy-3150-301-008-ed-3		
Document Title:	Statistical Analysis Plan Legacy 3150-301-008 Edition 3	
Document ID:	Doc ID-004424327	
Version Label:	3.0 CURRENT LATEST APPROVED	
Server Date (dd-MMM-yyyy HH:mm 'UTC'Z)	**Signed by**	**Meaning of Signature**
03-Apr-2023 09:28 UTC	PPD	Content Approval
03-Apr-2023 11:50 UTC	PPD PPD	Author Approval

Notes: (1) Document details as stored in ANGEL, an AstraZeneca document management system.