Statistical Analysis Plan

NCT Number: NCT04779307

Title: A Randomized, Double-Blind, Phase 3 Study to Evaluate the Efficacy and Safety of Vedolizumab Intravenous as Maintenance Therapy in Pediatric Subjects With Moderately to Severely Active Ulcerative Colitis Who Achieved Clinical Response Following Open-Label Vedolizumab Intravenous Therapy

Study Number: MLN0002-3024

Document Version and Date: Version 6, 05 June 2025

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

STATISTICAL ANALYSIS PLAN

Study Number: MLN0002-3024

Study Title: A Randomized, Double-Blind, Phase 3 Study to Evaluate the Efficacy and Safety of Vedolizumab Intravenous as Maintenance Therapy in Pediatric Subjects with Moderately to Severely Active Ulcerative Colitis Who Achieved Clinical Response Following Open−Label Vedolizumab Intravenous Therapy

Phase: 3

Version: 6.0

Date: June 5, 2025

Prepared by:

[redacted] Statistics GI

Based on:

Protocol Version: Amendment 7, Global

Protocol Date: 27 February 2025

Original SAP Date: 31 March 2021

REVISION HISTORY

Version of SAP	SAP Approval Date	Primary Rationale for Revision
Original 1.0	31 Mar 2021	Not Applicable
Version 2.0 Amendment 1.0	18 Feb 2022	To incorporate updates from Protocol Amendment 2
Version 3.0 Amendment 2.0	29 Nov 2023	To incorporate updates from Protocol Amendment 4
Version 4.0 Amendment 3.0	18 July 2024	To incorporate updates from Protocol Amendment 5
Version 5.0 Amendment 4.0	19 Nov 2024	To incorporate comments from Agency Based on Protocol Amendment 6
Version 6.0 Amendment 5.0	5 June 2025	To incorporate comments from Agency Based on Protocol Amendment 7

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

LIST OF IN-TEXT FIGURES

ABBREVIATIONS

5-ASA	5-aminosalicylates
6-MP	6-mercaptopurine
AE	adverse event
AESIs	AEs of special interest
ALT	alanine aminotransferase
AST	aspartate aminotransferase
AUC	area under the curve
AVA	anti-vedolizumab antibodies
AZA	azathioprine
bpm	beats per minutes
Cav	average serum concentration during a dosing interval
CI	confidence interval
Ctrough	serum trough concentration
CMV	cytomegalovirus
COVID-19	Coronavirus 2019
CRO	contract research organization
CRP	C-reactive protein
CSR	clinical study report
CV Units	Conventional Units
DSMB	data safety monitoring board
ECG	electrocardiogram
ET	Early Termination
eCRF	electronic case report form
FAS	full analysis set
FDA	Food and Drug Administration
FSH	follicle-stimulating hormone
GCDT	Global Clinical Development Team
GCP	Good Clinical Practice
GGT	gamma-glutamyl transferase
GI	gastrointestinal
HAHA	human antihuman antibody
hCG	human chorionic gonadotropin
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HLT	MedDRA high level term
IA	interim analysis
IAC	Independent Adjudication Committee
ICE	intercurrent event(s)
ICF	Informed Consent Form

ICH	International Conference on Harmonisation
ID	identification
IEC	independent ethics committee
INR	international normalized ratio
IRB	institutional review board
IRR	infusion-related reaction
IRT	interactive web response technology
IV	intravenous
LFT	liver function tests
LTFU	long-term follow-up
MedDRA	Medical Dictionary for Regulatory Activities
MES	Mayo Endoscopic Subscore
mMS	modified Mayo score
MTX	methotrexate
NOAEL	no-observed-adverse-effect level
NSAID	nonsteroidal anti-inflammatory drug
OTC	over-the-counter
PD	protocol deviation
PGA	Physician's Global Assessment
PK	pharmacokinetic
PML	progressive multifocal leukoencephalopathy
PT	preferred term
PTE	pretreatment event
PUCAI	Pediatric Ulcerative Colitis Activity Index
RBC	red blood cells
Q8W	once every 8 weeks
SAE	serious adverse event
SAP	statistical analysis plan
SI Units	International System of Units
SOC	MedDRA system organ class
SUSAR	suspected unexpected serious adverse reaction
TB	tuberculosis
TEAEs	treatment emergent AEs
TESAE	treatment emergent serious adverse event
TNF-α	tumor necrosis factor-alpha
UC	ulcerative colitis
ULN	upper limit of normal
VDZ	vedolizumab
WBC	white blood cell

1.0 OBJECTIVES, ENDPOINTS, ESTIMANDS

1.1 Objectives

1.1.1 Primary Objective

The primary objective of this study is to evaluate the efficacy of 2 different dose regimens of vedolizumab intravenous (IV) in pediatric subjects with moderately to severely active ulcerative colitis (UC) during maintenance therapy, based on clinical remission at Week 54.

1.1.2 Secondary Objectives

Secondary objectives of this study are to evaluate:

- The efficacy of high and low doses of vedolizumab IV in pediatric subjects with moderately to severely active UC, based on clinical remission at Week 14.
- The efficacy of high and low doses of vedolizumab IV as measured by sustained clinical remission at Weeks 14 and 54.
- The efficacy of high and low doses of vedolizumab IV as measured by sustained clinical response at Weeks 14 and 54.
- The efficacy of high and low doses of vedolizumab IV as measured by sustained endoscopic remission at Weeks 14 and 54.
- The efficacy of high and low doses of vedolizumab IV as measured by endoscopic response at Weeks 14 and 54.
- The effect of high and low doses of vedolizumab IV on achieving corticosteroid-free remission at Week 54.
- The effect of high and low doses of vedolizumab IV on clinical response over time up to Week 54.
- The efficacy of high and low doses of vedolizumab IV as measured by clinical remission over time up to Week 54.
- Vedolizumab pharmacokinetics (PK) in pediatric subjects with moderately to severely active UC after IV administration.
- Safety in pediatric subjects on maintenance therapy up to Week 54.
- The immunogenicity of vedolizumab in pediatric subjects with moderately to severely active UC treated with vedolizumab IV.
- The effect of vedolizumab on patterns of growth and pubertal development in pediatric subjects with moderately to severely active UC during their participation in the study.

1.1.3 Exploratory Objectives

Exploratory objectives of this study are to:

- Explore the relationship between vedolizumab exposure and the clinical and endoscopic response in pediatric subjects treated with vedolizumab IV during both induction and maintenance therapy.
- Assess markers of intestinal inflammation (C-reactive protein [CRP], serum albumin, fecal calprotectin) and analyze their correlation with trough PK levels of vedolizumab.
- Evaluate the quality of life in subjects aged 9 to 17 years who were treated with vedolizumab IV using the IMPACT-III questionnaire.
- Evaluate the efficacy of high and low doses of vedolizumab IV as measured by clinical remission based on Pediatric Ulcerative Colitis Activity Index (PUCAI).
- Evaluate the efficacy of high and low doses of vedolizumab IV as measured by clinical response based on PUCAI.

1.2 Endpoints

1.2.1 Primary Endpoint

The primary endpoint is clinical remission at Week 54, where clinical remission based on the modified Mayo score is defined as:

- Stool frequency subscore 0 to 1 and a decrease of 1 or more from baseline; and
- Rectal bleeding subscore of 0; and,
- Endoscopy subscore 0 to 1 (modified so that a score of 1 does not include friability).

1.2.2 Secondary Endpoints

Secondary endpoints include:

- Clinical remission at Week 14, where a subject achieves clinical remission if he or she meets the definition described in the primary endpoint.
- Sustained clinical remission at Week 54, where a subject achieves sustained clinical remission if he or she achieved clinical remission (as defined by primary endpoint) at Week 14 and at Week 54.
- Sustained endoscopic remission, defined as Mayo endoscopic subscore (MES) of ≤1 point, at Week 14 and at Week 54
- Endoscopic response, defined as a decrease from baseline in MES ≥1 point at Week 14
- Endoscopic response, defined as a decrease from baseline in MES ≥1 point at Week 54.
- Corticosteroid-free clinical remission at Week 54, where a subject achieves corticosteroid-free clinical remission at Week 54, if he or she meets the definition described in the primary endpoint and was off corticosteroids at least 12 weeks prior to and at Week 54.

- Clinical remission based on complete Mayo score at Week 54, where a subject achieves clinical remission if he or she achieved a complete Mayo score ≤2 points with no individual subscore >1 at Week 54.
- Serum trough concentrations of vedolizumab over time.
- Positive antivedolizumab antibodies (AVA)and positive neutralizing AVA during the study.
- Sustained clinical response of subjects at Weeks 14 and 54, where a subject meets clinical response if he or she has a reduction in complete Mayo score (see Section 9.4) of ≥3 points and ≥30% from baseline with an accompanying decrease in rectal bleeding subscore of ≥1 point or absolute rectal bleeding subscore of ≤1 point.
- Clinical response at Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54 where a subject achieves clinical response if he or she meets the following definition:
  - Reduction of ≥2 points and ≥25% from the baseline partial Mayo score, including a ≥1-point decrease in the Mayo stool frequency subscore and a ≥1-point reduction in the rectal bleeding subscore or absolute rectal bleeding subscore of ≤1 point.
- Clinical remission at Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54 where a subject achieves clinical remission based on partial Mayo score (a partial Mayo score of ≤2 points and no individual subscore >1 point).
- Safety assessments: descriptions of adverse events (AEs); SAEs; and AEs of special interest (AESIs), including evaluation of opportunistic infection, such as PML, liver injury, malignancies, infusion-related reactions, and hypersensitivity.
- Change from baseline in weight gain and linear growth z-score during the course of dosing with vedolizumab.
- Change in Tanner stage at Week 54 compared with baseline, each domain separately.

1.2.3 Exploratory Endpoints

Exploratory endpoints include:

- Change from baseline in CRP, albumin, and fecal calprotectin by visit.
- Change from baseline in IMPACT-III total and subscale scores at Weeks 14 and 54 for subjects aged 9 to 17 years at the time of the first dose.
- Change from baseline in Mayo and PUCAI scores and subscores by visit.
- Clinical remission based on PUCAI score at Week 54, where a subject achieves clinical remission if he or she has a PUCAI of <10 at Week 54.
- Clinical response based on PUCAI defined as a ≥20-point decrease from baseline in the PUCAI score for Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54.

1.3 Estimand(s)

The estimand framework for the primary estimand for the primary endpoint of clinical remission (based on the modified Mayo score) at Week 54 is summarized in Table 1.a.

Table 1.a Estimand Framework (Primary Estimand)

		Attributes			
Definition	**Treatment**	**A: Population**	**B: Variable or Endpoint**	**C: Strategy for Addressing Intercurrent Event (ICE)**	**D: Population-Level Summary**
Primary estimand of interest: Proportion of Week 54 clinical remitters for vedolizumab IV dose in maintenance phase, among pediatric subjects with moderately to severely active UC, meeting study eligibility criteria	Open-label Induction (Day 1, Week 2 and Week 6): Vedolizumab 300 mg IV (for subjects ≥30 kg) Vedolizumab 200 mg IV (for subjects >15 to <30) Vedolizumab 150 mg IV (for subjects 10 to 15 kg) Blinded Maintenance Treatment high or low dose (every 8 weeks; Week 14 to Week 46) Vedolizumab 300 mg or 150 mg IV (for subjects ≥30 kg) Vedolizumab 200 mg or 100 mg IV (for subjects >15 to <30 kg) Vedolizumab	Pediatric subjects with moderately to severely active UC, meeting study eligibility criteria. ITT-M Analysis Set	Clinical remission (based on modified Mayo score) at Week 54	ICEs: • Discontinuation of treatment during maintenance phase. • Dose escalation from low dose to high dose during the maintenance phase. • Receiving medication(s) not allowed by the protocol. Refer to 1.3.1 for details. <u>Strategy:</u> Composite strategy will be used for addressing ICEs; subjects with above ICEs will be treated as non-remitters	Proportion of Week 54 clinical remitters for vedolizumab IV dose in maintenance phase

Table 1.a Estimand Framework (Primary Estimand)

		Attributes			
Definition	**Treatment**	**A: Population**	**B: Variable or Endpoint**	**C: Strategy for Addressing Intercurrent Event (ICE)**	**D: Population-Level Summary**
	150 mg or 100 mg IV (for subjects 10 to 15 kg)				

[1] Prohibited and some permitted medications are considered ICE due to potential impact on disease under study (Ulcerative Colitis).

Similar secondary estimands will be analyzed for the following secondary efficacy endpoints at Week 54.

- Sustained clinical remission at Week 14 and Week 54
- Sustained endoscopic remission at Week 14 and at Week 54
- Endoscopic response at Week 54
- Corticosteroid-free clinical remission at Week 54
- Clinical remission based on complete Mayo score at Week 54.
- Sustained clinical response at Week 14 and 54

1.3.1 ICEs involving AxMP or Other Medications

Receiving the medications defined in the sections below will be considered intercurrent events (ICE).

1.3.1.1 Corticosteroids

Subjects starting rescue IV or oral corticosteroids during the maintenance phase will be considered non-responders. The corticosteroids for UC or UC equivalent indication will be considered a rescue, if any of the following conditions are met.

- If the steroid started on or after first maintenance dose of the study drug, or
- the route of corticosteroid (started before the first maintenance dose of the study drug) was changed from oral to IV, or
- the prednisone equivalent total daily dose (started before the first maintenance dose of study drug) is increasing during maintenance, or
- the steroid use is not continuous. The continuous is based on the frequency of the medication. Therefore, a medication of once a week could have a gap of 7 days, whereas a medication of biweekly could have a gap of 14 days.

However, in subjects who were on steroids at baseline, tapering of steroids during maintenance will not be considered rescue. To determine whether the corticosteroid is tapering, medications will be identified based on Auxiliary Medication with the ATC Code of A07EA and H02AB, which will be converted to a prednisone equivalent total daily dose, and decreasing total daily dose will be considered tapering.

1.3.1.2 Prohibited Medications

Subjects taking prohibited medications will be considered as having an ICE. The following will be considered as prohibited medications:

- Any "TUMOR NECROSIS FACTOR ALPHA (TNF-) INHIBITORS" taken between 14 days prior to Day 1 and Week 54/ET
- Any "SELECTIVE IMMUNOSUPPRESSANTS" or "INTERLEUKIN INHIBITORS" or "CALCINEURIN INHIBITORS" taken between informed consent date and Week 54/ET
- Any "OTHER ANTIINFLAMMATORY AND ANTIRHEUMATIC AGENTS, NON-STEROIDS" taken between Day 1 and Week 54/ET with at least 2 consecutive weeks of use
- Any "AMINOSALICYLIC ACID AND SIMILAR AGENTS" or "CORTICOSTEROIDS ACTING LOCALLY" by topical or rectal route taken between Day – 14 and Week 54/ET

1.3.1.3 Permitted Medications

The following medications are permitted per protocol at a stable dose as described below.

- "OTHER IMMUNOSUPPRESSANTS" must be at stable dose for at least 8 weeks prior to first dose of study drug and throughout the first 14 weeks of the study
- Oral 5-ASA under the ATC Class "AMINOSALICYLIC ACID AND SIMILAR AGENTS" for UC and UC equivalent indication must be stable for at least 2 weeks prior to first dose of study drug and throughout the first 14 weeks of the study. The following medications are considered Oral 5-ASA: Mesalamine/Mesalazine, Olsalazine, Sulfasalazine and Balsalazide

However, any increase in total daily dose or initiation of these medications will be considered an ICE as per below:

- "OTHER IMMUNOSUPPRESSANTS" increase in total daily dose or initiation between 8 weeks prior to Day 1 and Week 54/ET
- Oral 5-ASA under the ATC Class "AMINOSALICYLIC ACID AND SIMILAR AGENTS" for UC and UC equivalent indication increase in total daily dose or initiation between 2 weeks prior to Day 1 and Week 54/ET. The following medications are considered oral 5-ASA: Mesalamine/Mesalazine, Olsalazine, Sulfasalazine and Balsalazide

Other unstable permitted medications such as probiotics or antidiarrheal drugs will be considered as protocol deviations (PDs), and not an ICE.

2.0 STUDY DESIGN

This is a phase 3 study of vedolizumab administered during a 14-week, open–label induction period followed by a 40-week, 2-dose arm, randomized, double-blind maintenance period designed to evaluate the efficacy, safety, PK, and immunogenicity of high and low doses in 3 different weight groups of vedolizumab IV for maintenance of remission in children aged 2 to 17 years, inclusive, who weigh ≥10 kg with moderately to severely active UC. The data from this study will be used for partial extrapolation of efficacy from adults to pediatric subjects based on the efficacy, safety, and PK data from completed studies in adult subjects.

Screening: *Subjects must have been diagnosed with moderately to severely active UC at least 1 month before screening and failed response to, lost response to, or been intolerant to at least 1 of the current standard of care induction and maintenance therapies for UC, including immunomodulators (e.g., AZA, 6-MP, methotrexate [MTX]), and TNF-α antagonists (eg, infliximab, adalimumab). Lack of response to corticosteroids at the time of the initial diagnosis of UC or inability to wean off corticosteroids without a recurrence of symptoms will also qualify a subject for the study. Subjects or their caregivers will be provided with an electronic symptom diary and trained on its use during the initial screening visit and will complete the symptom diary daily throughout participation in the study. Subjects will be evaluated during a 35-day screening period by using symptom diaries, colonoscopy, and determination of fecal calprotectin and CRP. Those with moderately to severely active UC, defined by a modified Mayo score of 5 to 9 (sum of the stool frequency, rectal bleeding, and endoscopic subscores with an endoscopic subscore of ≥2) (and meeting all other eligibility criteria), can enter the open–label induction period. The PUCAI will also be assessed. All colonoscopies (at screening and all subsequent flexible sigmoidoscopies at Weeks 14 and 54 or the ET visit) will be centrally read for Mayo endoscopic subscore. Historical colonoscopy (with recorded video) done up to 30 days prior to signing informed consent, can be submitted to central readers for assessment. If found to be adequate by central readers, this video colonoscopy will be used for the assessment of Mayo endoscopic subscore, which can be used as the baseline assessment.*

Induction Period: *On Day 1, the modified Mayo score should be verified and documented before infusion. During the induction period, subjects ≥30 kg will receive open–label vedolizumab, 300 mg IV at Day 1, Week 2, and Week 6; subjects who weigh >15 to <30 kg will receive an open–label dose of vedolizumab 200 mg IV at Day 1, Week 2, and Week 6; and subjects who weigh 10 to 15 kg will receive an open–label dose of vedolizumab 150 mg IV at Day 1, Week 2, and Week 6. These doses were determined using an established population PK model, developed using data from adult phase 3 studies and the pediatric phase 2 study (MLN0002-2003) for ≥30 kg and <30 kg cohorts, to match the vedolizumab exposure seen in the adult populations during the induction phase. Approximately 120 subjects will be enrolled in the induction period with around three-fourths of subjects weighing ≥30 kg and around one-fourth of subjects weighing 10 to <30 kg at the time of enrollment into the study.*

Weight Group [a]	*Induction Dose*
Subjects ≥30 kg	*300 mg IV at Day 1, Week 2 and Week 6*
Subjects >15 to <30 kg	*200 mg IV at Day 1, Week 2 and Week 6*
Subjects 10 to 15 kg	*150 mg IV at Day 1, Week 2 and Week 6*

IV: intravenous.

[a] No change in induction dose for weight change over time.

At Week 14, all subjects will be evaluated for clinical response, based on modified Mayo score (defined as a reduction of ≥2 points and ≥30% from the baseline modified Mayo score with an accompanying decrease in rectal bleeding subscore of ≥1 point or absolute rectal bleeding subscore of ≤1 point). Those who do not achieve a clinical response during the open–label induction period will discontinue from the study, complete the end-of-study (EOS)/early termination (ET) visit, and proceed to the follow-up safety assessment 18 weeks after their last dose of study drug. These subjects will then enter Study MLN0002-3029 for an observational long-term follow-up (LTFU) period of 2 years after the last dose of study drug. During the LTFU period, data will be collected either by clinic visit OR, if site attendance is not feasible, by phone call every 6 months.

Randomization Into the Maintenance Period: *At Week 14, those who achieve clinical response (defined as a reduction of ≥2 points and ≥30% from the baseline modified Mayo score with an accompanying decrease in rectal bleeding subscore of ≥1 point or absolute rectal bleeding subscore of ≤1 point) will be stratified by previous exposure/failure to TNF-α antagonists therapy or naive to TNF-α antagonists therapy, and by weight group (≥30 kg; >15 kg to <30 kg; 10 kg to 15 kg). Subjects will be randomized 1:1 to one of the 2 dose groups (high dose, low dose), within each of the 3 weight groups (≥30 kg; >15 to <30 kg; 10 to 15 kg), to one of the 2 dose groups (high dose, low dose), for the double-blind maintenance period.*

Weight Group [a]	*High Dose*	*Low Dose*
Subjects ≥30 kg	*300 mg IV Q8W*	*150 mg IV Q8W*
Subjects >15 to <30 kg	*200 mg IV Q8W*	*100 mg IV Q8W*
Subjects 10 to 15 kg	*150 mg IV Q8W*	*100 mg IV Q8W*

IV: intravenous; Q8W: once every 8 weeks.

[a] During the maintenance period, adjustments to the dose will be as described for lack of maintenance of clinical response and worsening of disease only; no adjustments made strictly on the basis of weight change over time.

Maintenance Period: *Subjects who qualify as responders at the end of the induction period (Week 14) will initiate blinded IV doses of vedolizumab Q8W as maintenance dose through Week 46. The maintenance doses for the 3 weight groups were chosen similar to the way induction doses were chosen. A population PK model was used to match the vedolizumab exposure seen in phase 3 studies for the adult population during the maintenance phase. The high and low doses were selected to maximize the dose difference between the two arms while maintaining the exposure within the adult population range. Based on the chosen induction dose, a high and low maintenance dose of 300 mg and 150 mg was chosen for the ≥30 kg weight group, 200 mg and 100 mg was chosen for >15 to <30 kg weight group, and 150 mg and 100 mg was chosen for the 10 to 15 kg weight group. A 150 mg high maintenance dose in the 10 to 15 kg weight group was predicted to result in an exposure closer to, but not exceeding, the upper limit of the adult*

exposure in the maintenance phase. Similarly, a 100 mg low maintenance dose was predicted to result in an exposure closer to, but not below, the lower limit of adult exposure in the maintenance phase. A 100 mg low maintenance dose was predicted to result in similar and consistent exposure among the entire <30 kg weight groups. A dose of 75 mg was considered for the low maintenance dose in the 10 to 15 kg weight cohort; however, this was predicted to have resulted in an exposure less than adult phase 3 exposure.

At each visit before the infusion, clinical response based on partial Mayo score will be assessed. Improvement or maintenance of clinical response will result in continuation of the blinded dosing of vedolizumab IV. Lack of maintenance of clinical response and worsening of disease is defined as the following at 2 consecutive visits (scheduled or unscheduled) at least 7 days apart:

- *Increase in the Mayo subscore of stool frequency by 1 point to at least ≥2 points; and*
- *Increase in the Mayo subscore of rectal bleeding by at least 1 point with an absolute bleeding subscore of ≥2 points.*

In addition to safety assessment and evaluation of symptoms and laboratory parameters indicative of disease activity at each visit, an endoscopy will be performed at Week 54 to evaluate endoscopic response and remission. Blood samples will be collected for vedolizumab PK and immunogenicity evaluation throughout the study. Subjects who continue to maintain a corticosteroid-free clinical response at Week 54 will be eligible to receive study drug in the MLN0002-3029 extension study.

For subjects who do not maintain corticosteroid-free clinical response at Week 54, or who discontinue study drug at any time during the induction or maintenance periods of this study, final on-study assessments will include an EOS or ET visit and a follow-up safety visit 18 weeks after the last dose of study drug. Subjects will then enter Study MLN0002-3029 for an observational LTFU period of 2 years after the last dose of study drug. During the LTFU period, data will be collected either by clinic visit OR, if site attendance is not feasible, by phone call every 6 months.

Dose Escalation and Steroid Rescue During the Maintenance Period: *Lack of maintenance of clinical response and worsening of disease will result in a blinded dose escalation based on the weight at the time of the worsening of disease as described in the table below.*

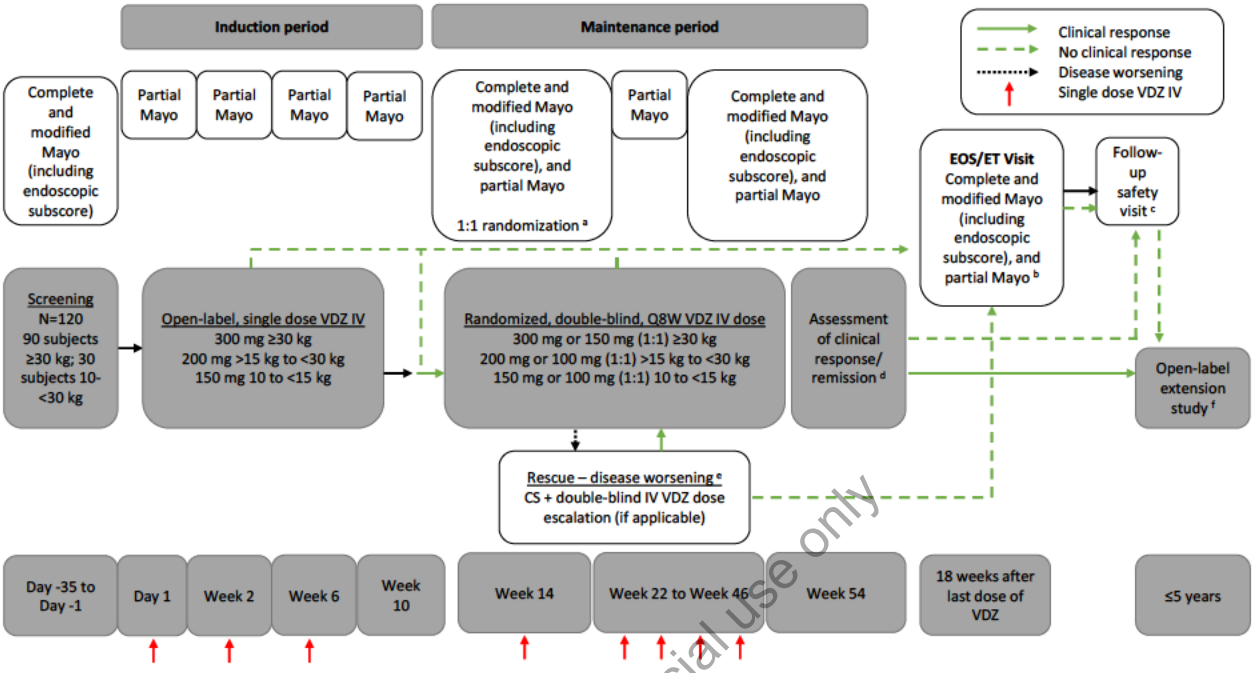
Weight Group (at Time of Randomization Into the Maintenance Period)	*Blinded Dose (at Randomization)*	*Weight Group (at Time of Disease Worsening)*	*Dose Escalation?*	*Blinded Dose (After Disease Worsening)*
10 to 15 kg	*100 mg IV Q8W*	*10 to 15 kg*	*Yes*	*150 mg IV Q8W*
		>15 to <30 kg	*Yes*	*200 mg IV Q8W*
		≥30 kg	*Yes*	*300 mg IV Q8W*
	150 mg IV Q8W	*10 to 15 kg*	*No*	*150 mg IV Q8W*
		>15 to <30 kg	*Yes*	*200 mg IV Q8W*
		≥30 kg	*Yes*	*300 mg IV Q8W*
>15 to <30 kg	*100 mg IV Q8W*	*10 to 15 kg*	*Yes*	*150 mg IV Q8W*
		>15 to <30 kg	*Yes*	*200 mg IV Q8W*
		≥30 kg	*Yes*	*300 mg IV Q8W*
	200 mg IV Q8W	*10 to 15 kg*	*No*	*200 mg IV Q8W*
		>15 to <30 kg	*No*	*200 mg IV Q8W*
		≥30 kg	*Yes*	*300 mg IV Q8W*
≥30 kg	*150 mg IV Q8W*	*10 to 15 kg*	*No*	*150 mg IV Q8W*
		>15 to <30 kg	*Yes*	*200 mg IV Q8W*
		≥30 kg	*Yes*	*300 mg IV Q8W*
	300 mg IV Q8W	*10 to 15 kg*	*No*	*300 mg IV Q8W*
		>15 to <30 kg	*No*	*300 mg IV Q8W*
		≥30 kg	*No*	*300 mg IV Q8W*

IV: intravenous; Q8W: once every 8 weeks.

Lack of maintenance of clinical response and worsening of disease will also result in subjects who are receiving vedolizumab IV to become eligible to receive rescue therapy with corticosteroids at investigator discretion in an attempt to re-establish response/remission during the maintenance period. Those subjects whose dose is escalated and/or who receive rescue corticosteroids will, for the analysis of blinded maintenance dose groups, be considered nonresponders to the initial blinded dose of vedolizumab; however, they can remain in the study so safety, PK, and efficacy data will continue to be assessed in this cohort for the duration of the study. If the investigator elects to continue the subject in the study and initiates corticosteroid rescue at a scheduled or unscheduled visit, it must be based on evaluation of Mayo score with documentation of worsening. Prescription for a dose of corticosteroids at the discretion of their physician should not exceed 40 mg or 1 mg/kg (whichever is lower) of oral prednisone per day or its equivalent. All subjects, including those who are initiated on corticosteroid rescue therapy, who do not achieve clinical response in 2 subsequent consecutive visits should be discontinued from the study and proceed to the EOS/ET visit, followed by the follow-up safety visit 18 weeks after their last dose of study drug. These subjects will then enter Study MLN0002-3029 for an observational LTFU period of 2 years after the last dose of study drug. During the LTFU period, data will be collected either by clinic visit OR, if site attendance is not feasible, by phone call every 6 months.

After initiation of corticosteroid rescue therapy, tapering of corticosteroids will start if/when the subject has achieved clinical response based on Mayo score measured at the next scheduled or unscheduled follow-up visit. Corticosteroids will be tapered on a controlled schedule defined in the protocol, designed to wean subjects off corticosteroids within 12 weeks of initiation of tapering. Worsening of disease on tapering of corticosteroids will result in discontinuation from the study; the subject will proceed to the EOS/ET visit followed by the follow-up safety visit 18weeks after their last dose of study drug. These subjects will enter study MLN0002-3029 for an observational LTFU period of 2 years after the last dose of study drug. During the LTFU period, data will be collected either by clinic visit OR, if site attendance is not feasible, by phone call every 6 months.

Figure 2.a Schematic of Study Design



CS: corticosteroid; EOS: end of study; ET: early termination; IV: intravenous; LTFU: long-term follow-up; PK: pharmacokinetics; UC: ulcerative colitis; VDZ: vedolizumab.

[a] At Week 14, subjects who achieve a clinical response will be randomized 1:1 to a double-blind maintenance period starting at Week 14 through Week 54. If no clinical response is achieved by Week 14, the subject will discontinue from the study, proceed to the EOS/ET visit, and a follow-up safety visit 18 weeks after the last dose of study drug and will continue into the extension study for an observational LTFU period of 2 years after the last dose of study drug. During the LTFU period, data will be collected either by clinic visit OR, if site attendance is not feasible, by phone call every 6 months.

[b] If discontinuation during the maintenance phase occurs before or at the time of Week 30 visit, the endoscopy required for the ET visit will be waived.

[c] Final assessment in this study will include a follow-up safety visit 18 weeks after the last dose of study drug for subjects who do not continue into the extension study for continued treatment with vedolizumab and for subjects who discontinue study drug at any time during the study. These subjects will continue into the extension study for an observational LTFU period of 2 years after the last dose of study drug. During the LTFU period, data will be collected either by clinic visit OR, if site attendance is not feasible, by phone call every 6 months.

[d] At Week 54, subjects will undergo flexible sigmoidoscopy in addition to symptom, safety, and laboratory evaluation to fully assess clinical and endoscopic response/remission.

[e] Failure to maintain clinical response and worsening of disease will result in the eligibility of subjects who are receiving vedolizumab IV to receive additional rescue therapy with corticosteroids at investigator discretion in an attempt to re-establish response/remission during the maintenance period. These subjects will be considered nonresponders/nonremitters to the initial blinded dose of vedolizumab; however, safety, PK, and efficacy data will continue to be assessed.

[f] Subjects who continue to maintain a corticosteroid-free clinical response at Week 54 may continue treatment with vedolizumab in the extension study for up to 5 years, or until approval of vedolizumab for children, or cessation of vedolizumab drug development for UC in children.

Table 2.a Schedule of Study Procedures

	Screening		Induction Period				Maintenance Period					Follow-up Safety Visit [a]	Unscheduled Visit [b]
	Days -35 to -1	Day 1	Wk 2	Wk 6	Wk 10	Wk 14	Wk 22	Wk 30	Wk 38	Wk 46	Wk 54/ EOS/ET Visit [a]		
	NA	NA	±2 days	±2 days	±3 days	±7 days	±3 days	±3 days	±3 days	±3 days	±7 days	±1 week	
Assessment	**Visit 1**	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	**NA**	
Parental informed consent/subject assent [c]	X	X	X	X	X	X	X	X	X	X	X	X	X
Allergy reaction card	X												
Patient alert card	X												
Demographics	X												
Medical history	X												
Medication history	X												
Tanner staging	X										X		
Physical examination [d]	X	X	X	X	X	X	X	X	X	X	X	X	X
Weight and height [e]	X	X	X	X	X	X	X	X	X	X	X	X	X
Vital signs [e]	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant medications	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant procedures	X	X	X	X	X	X	X	X	X	X	X	X	X
IRT accessed	X	X	X	X		X	X	X	X	X	X		
PTE assessment [f]	X	X											
AE assessment [g]		X	X	X	X	X	X	X	X	X	X	X	X
Endoscopy [h]	X					X					X		X
Dosing		X	X	X		X	X	X	X	X			
Clinical Outcomes Assessments													
PUCAI	X	X	X	X	X	X	X	X	X	X	X		X
Complete Mayo	X					X					X		X
Modified Mayo	X	X				X					X		X
Partial Mayo		X	X	X	X	X	X	X	X	X	X		X

Table 2.a Schedule of Study Procedures

	Screening	Induction Period				Maintenance Period						Follow-up Safety Visit [a]	Unscheduled Visit [b]
	Days -35 to -1	Day 1	Wk 2	Wk 6	Wk 10	Wk 14	Wk 22	Wk 30	Wk 38	Wk 46	Wk 54/ EOS/ET Visit [a]		
	NA	NA	±2 days	±2 days	±3 days	±7 days	±3 days	±3 days	±3 days	±3 days	±7 days	±1 week	
Assessment	**Visit 1**	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	**NA**	
Electronic diary [i]	X	X	X	X	X	X	X	X	X	X	X		X
IMPACT-III		X				X					X		
Laboratory assessments													
TB screening [j]	X												
Serum pregnancy test [k]	X										X	X	
Urine pregnancy test [k]		X	X	X	X	X	X	X	X	X			X
HBV, HCV, HIV	X												
Stool samples (culture, O&P, *C diff*) [l]	X												X
Stool sample for fecal calprotectin [m]	X					X					X		X
Clinical chemistry (incl CRP)	X	X	X	X	X	X	X	X	X	X	X	X	X
Hematology	X	X	X	X	X	X	X	X	X	X	X	X	X
PK assessment [n]		X	X	X	X	X	X	X	X	X	X	X	X
AVA assessment [o]		X		X		X		X		X	X	X	X

AE: adverse event; AVA: antivedolizumab antibody; CD: Crohn's disease; *C diff*: *Clostridioides difficile*; CRP: C-reactive protein; EOS: end-of-study; ET: early termination; HBV: hepatitis B virus; HCV: hepatitis C virus; incl: including; IRT: interactive response technology; NA: not applicable; O&P: ova and parasite; PK: pharmacokinetic; PTE: pretreatment event; PUCAI: Pediatric Ulcerative Colitis Activity Index; SAE: serious adverse event; TB: tuberculosis; UC: ulcerative colitis; Wk: week.

[a] EOS/ET and follow-up safety visits: After the Week 54 EOS/ET visit assessments have been completed, subjects may be eligible to receive continued treatment with vedolizumab in extension Study MLN0002-3029. Subjects who do **not** qualify to receive continued treatment in the extension study or subjects who discontinue from the study for any reason will complete the EOS/ET visit and the follow-up safety visit (18 weeks after last dose). These subjects will then enter Study MLN0002-3029 for an observational LTFU period of 2 years after the last dose of study drug. In case of ET at the time of a scheduled visit, any assessments for ET already performed as part of the scheduled visit will not be repeated. If discontinuation occurs within 2 weeks after a scheduled visit; only assessments pertaining to the EOS visit (and not completed in the most recent scheduled visit) will be required. The investigator may repeat any of the scheduled visit assessments as deemed clinically appropriate.

[b] Subjects seen at an unscheduled visit for disease exacerbation will complete the unscheduled visit assessments. Other unscheduled visits may include concomitant medications, vital signs, clinical laboratory blood draws, PK assessment, AVA assessment, urine sample, and AE collection as appropriate. IRT may include an unscheduled visit to allow for sites to make dose modifications between scheduled visits or to place a subject on dose hold. Endoscopy may be performed (only if deemed necessary per the investigator's discretion) AND complete/modified Mayo scores may be obtained (only if an endoscopy is performed).

Table 2.a Schedule of Study Procedures

	Screening		Induction Period				Maintenance Period					Follow-up Safety Visit [a]	Unscheduled Visit [b]
	Days -35 to -1	Day 1	Wk 2	Wk 6	Wk 10	Wk 14	Wk 22	Wk 30	Wk 38	Wk 46	Wk 54/ EOS/ET Visit [a]		
	NA	NA	±2 days	±2 days	±3 days	±7 days	±3 days	±3 days	±3 days	±3 days	±7 days	±1 week	
Assessment	**Visit 1**	**Visit 2**	**Visit 3**	**Visit 4**	**Visit 5**	**Visit 6**	**Visit 7**	**Visit 8**	**Visit 9**	**Visit 10**	**Visit 11**	**NA**	

[c] Subject age will be assessed at each study visit to confirm that their current informed consent/pediatric assent is applicable. If informed consent/pediatric assent needs to be updated because of age change, this will be done before any study procedures.
[d] Physical examination: Clinically significant findings will be recorded as AEs if they start after the first dose of study drug.
[e] Weight, height, and vital signs will be measured on clinic dosing days before dosing.
[f] PTEs will be captured immediately following the signing of the informed consent/pediatric assent at screening and until the first dose of study drug.
[g] Collection of AEs, to include concomitant medications, and SAEs will begin following the first dose of study drug and will be done before each dose on dosing days through the Week 54/EOS/ET visit and follow-up safety visit, if applicable, and at any unscheduled visit.
[h] Colonoscopy will be performed at screening within 30 days of signing informed consent/pediatric assent. Video endoscopy (flexible sigmoidoscopy) to be performed at Week 14 and Week 54, or EOS/ET if subject discontinues study drug early during the maintenance phase (after Week 30). When scheduling the endoscopy, allow sufficient time for centrally read results as well as assessment of laboratory results and severity of symptoms. Because the endoscopy preparation can interfere with the assessment of other clinical parameters, diary entries used to calculate the complete Mayo score should not be from days during which the endoscopy preparation is administered.
[i] Electronic diary training will be provided to subjects/parents/legal guardians at screening. Diary entries are to be completed daily for all subjects. An average of 3 days of completed diary data before the Day 1 visit will be required for Mayo score calculation. Diary entries preceding each subsequent study visit (7 days) will be used to calculate the Mayo score. Because the endoscopy preparation can interfere with the assessment of other clinical parameters, diary entries used to calculate the complete Mayo score should not be from days during which the endoscopy preparation is administered.
[j] TB screening may be completed using QuantiFERON or the tuberculin skin test.
[k] All female subjects who are menstruating or aged ≥11 years, whichever is younger, must have a serum pregnancy test at screening and at the Week 54/EOS/ET visit and the follow-up safety visit. A urine pregnancy test will be completed before each dose of study drug. Urine pregnancy tests may be performed locally.
[l] A stool sample will be obtained for culture, ova and parasite evaluation, and *C difficile* assay. A sample will be collected and cultured during screening and at any point in the study when a subject becomes symptomatic, including worsening of UC. Stool tests may be performed locally at discretion of the investigator.
[m] Fecal calprotectin stool sample must be collected before any bowel preparation that is given for endoscopy. These samples must be analyzed by the central laboratory.
[n] Day 1: 1 sample (obtained postdose); Weeks 2, 6, and 46: 2 samples (predose and postdose); Weeks 14, 22, 30, and 38: 1 sample (predose); Week 10, Week 54/EOS/ET, final safety visit, and any unscheduled visit: 1 sample (any time during visit). Predose PK samples will be obtained within 30 minutes before dosing. Postdose PK samples will be obtained within 60 minutes after the end of the infusion.
[o] All AVA samples will be obtained predose (within 30 minutes before dosing) if collected on dosing days.

3.0 STATISTICAL HYPOTHESES AND DECISION RULES

3.1 Statistical Hypotheses

Not applicable.

3.2 Statistical Decision Rules

Not applicable.

3.3 Multiplicity Adjustment

Not applicable.

4.0 SAMPLE-SIZE DETERMINATION

The primary statistical objective is the estimation of the primary endpoint, clinical remission at Week 54, by dose groups (high dose, low dose). The sample-size justification was based on ensuring adequate precision (based on the half-width of the 95% CI) for the estimate of the clinical remission rate.

Assuming the true clinical remission rate of 42% at Week 54, a sample size of 36 subjects in each maintenance arm will provide a 95% Clopper-Pearson exact CI with a lower bound of 25.5% and an upper bound of 59.2%. In addition, the table below summarizes the Clopper-Pearson exact 95% CI for the clinical remission rates ranging from 37% to 47% with a sample size of 36 subjects. The width of this CI provides adequate precision around the primary endpoint of interest.

Observed Clinical Remission Rates With Clopper-Pearson 95% CI (for N=36 subjects)	
Observed Remission Rate	***95% Clopper-Pearson CI***
37%	*(20.8%-53.8%)*
42%	*(25.5%-59.2%)*
47%	*(27.9%-61.9%)*

To ensure a randomized sample size of 72 subjects (36 per maintenance arm) during the maintenance phase, approximately 120 subjects will need to be enrolled in the induction phase of the study.

5.0 ANALYSIS SETS

5.1 Safety Analysis Set

The Safety Analysis Set (SAF) will include all subjects who receive at least 1 dose of study vedolizumab. Subjects in this set will be analyzed according to the treatment actually received.

The Safety Analysis Set – Induction (SAF-I) will include all subjects who receive at least 1 induction dose but were not randomized to maintenance phase.

The Safety Analysis Set – Maintenance (SAF-M) will include all subjects who receive at least 1 maintenance dose.

The SAF, and SAF-I will be analyzed according to the first actual weight/dose level subject received in the induction phase. The SAF-M will be analyzed according to the first actual dose level subject received in the maintenance phase (low/high/overall).

5.2 Intent-to-Treat

Intent-to-treat (ITT): The ITT analysis set will include all enrolled subjects for the induction phase, defined by when informed consent form is obtained and not a screen failure. Subjects are not required to have received any dose of vedolizumab to be included in the ITT.

Intent-to-treat – Maintenance (ITT-M) will include all subjects randomized in the maintenance phase.

Subjects in these analysis sets are not required to receive any dose of vedolizumab to be included. Subjects will be analyzed according to the treatment they were intended (induction phase) or randomized (maintenance phase) to receive, regardless of any errors of dosing.

The ITT (weight/dose groups) and ITT-M (low/high/overall) will be used for efficacy analysis.

5.3 Modified Intent-to-Treat

Modified Intent-to-treat (mITT): The mITT will include all randomized subjects who receive at least 1 dose of study drug during the maintenance. Subjects who only receive induction therapy and not randomized into the maintenance phase will not be included in the mITT. Subjects in this set will be analyzed according to the treatment they were randomized to receive. If mITT is same as ITT-M, the mITT tables will not be produced.

5.4 Per-Protocol Analysis Set

Per-protocol Set – Maintenance (PPS-M): The PPS-M is a subset of the ITT-M. The PPS-M consists of all subjects who do not violate the terms of the protocol in a way that would impact the study outcome significantly.

The reasons to exclude subjects from PPS-M will include the following:

a) Subject does not have confirmed diagnosis of UC
    a. inclusion criteria number 3: The subject has moderately to severely active UC, unresponsive or intolerant to their current SOC
    b. inclusion criteria number 6: Subjects with UC diagnosed at least 1 month before screening. Subjects with moderately to severely active UC based on a modified Mayo score of 5 to 9 with a Mayo endoscopic subscore of $\geq$2 at screening endoscopy
    c. inclusion criteria number 8: Subjects with evidence of UC extending proximal to the rectum at a minimum

d. exclusion criteria 9: Subjects who have had subtotal or total colectomy or have a jejunostomy, ileostomy, colostomy, ileo-anal pouch, or known fixed stenosis of the intestine

e. exclusion criteria 11: Subjects with a current diagnosis of indeterminate colitis.

f. exclusion criteria 12: Subjects with clinical features suggesting monogenic very early onset IBD.

b) Subject does not have study drug compliance (>= 80%) based on study participation

c) Subject did not receive correct study drug (i.e., induction phase, doses = weight-based doses, and at maintenance phase, randomized doses = actual doses received)

d) Subject does not remain blinded in the maintenance phase through Week 54

e) Subject does not have confirmed modified Mayo score (between 5-9) at baseline based on Day 1 PGA Score

f) Subject was randomized without a clinical response at Week 14, defined as a reduction of ≥2 points and ≥30% from the baseline modified Mayo score with an accompanying decrease in rectal bleeding subscore of ≥1 point or absolute rectal bleeding subscore of ≤1 point)

g) Subject does not have a PGA at Week 14 between the analysis day of 85 and 113

h) Subject does not have a PGA at Week 54 between the analysis day of 365 and 393

i) Subject had prohibited medication during the study

j) Subject received at least one permitted medication not according to the protocol

k) Subject did not follow proper washout for prior biologic.

a. exclusion criteria 4: Subjects who have received either (1) an investigational biologic or (2) an approved biologic or biosimilar agent.

All protocol deviations will be reviewed to assess potential impact on study results; additional deviations leading to exclusion from the PPS-M will be documented in the CSR. The review and assessment of subject's eligibility for the PPS-M will be done prior to database lock.

All deviations leading to exclusion from the PPS-M will be summarized and listed.

Analyses of primary efficacy endpoints will be performed using the PPS-M as a sensitivity analysis, if more than 5% of the total subjects in the ITT are excluded from the PPS-M.

5.5 Pharmacokinetic Analysis Set

Pharmacokinetic (PKAS) Analysis Set: The PK evaluable population is defined as all subjects who receive at least 1 dose of vedolizumab and have measurable drug concentration in blood to allow for PK evaluation.

Pharmacokinetic Analysis Set - Induction (PKAS-I) is all subjects who receive at least 1 induction dose with measurable drug concentration in blood to allow for PK evaluation but were not randomized into the maintenance phase.

Pharmacokinetic Analysis Set – Maintenance (PKAS-M) is all subjects who receive at least 1 maintenance dose with measurable drug concentration in blood to allow for PK evaluation.

6.0 STATISTICAL ANALYSIS

6.1 General Considerations

Baseline value is defined as the last observed non-missing value before the first dose of study drug (see Appendix Section 9.2.2).

A visit windowing convention will be used to determine the analysis value for a given study visit for observed data analyses (see Appendix Section 9.2.4 for details).

The analysis based on ITT, SAF, and PKAS-I analysis set would be based on weight cohort according to participant's weight at baseline. The analysis based on ITT-M, SAF-M, mITT, PPS-M and PKAS-M would be based on weight and dose group at beginning of maintenance.

6.1.1 Handling of Treatment Misallocations

Subjects who received treatment (dose) other than allocated as per randomization will be handled according to the definitions of the respective analysis sets.

6.1.2 Analysis Approach for Continuous Variables

Descriptive statistics to be presented for the continuous variables. Number of subjects, mean, median and quartiles will be presented to 1 more decimal place than the recorded data. The standard deviations (SDs) and quartiles for PK data will be presented to 2 more decimal places than the recorded data. The minimum and maximum will be presented with the same decimal places as the recorded data.

CIs will be displayed to the same number of decimal places as the point estimate.

Where applicable, change from baseline will be derived as post-baseline value – baseline value.

6.1.3 Analysis Approach for Binary Variables

Categorical data will be summarized using the number and percentage of subjects for each category where appropriate. All percentages will be presented with 1 decimal place.

For the binary endpoints in the induction phase the number and percentage of subjects achieving the respective endpoint will be summarized.

For the binary response-type efficacy endpoints, the 95% Confidence Interval (CI) will be calculated based on the Clopper Pearson method, unless otherwise specified. The 95% CI will be displayed for both Yes (responder) and No (non-responder).

For the binary endpoints of the induction phase and maintenance, the number, percentage of subjects achieving the respective endpoint, and the 95% CI will be presented.

For all binary efficacy endpoints in maintenance phase, the absolute difference between the high dose and low dose and its associated 95% CI, based on the Wald method, by dose group (and overall) will be presented.

For the primary endpoint and secondary efficacy binary endpoints at Week 54 for the ITT-M, mITT, and PPS-M analysis sets, the following will be presented:

- the common Cochran-Mantel-Haenszel (CMH) risk difference and the associated 95% CI, stratified by weight group at Week 14 and previous TNF-alpha antagonist therapy status based on IRT data.
- the CMH risk difference and the associated 95% CI, stratified by weight group, for each TNF-alpha antagonist therapy status stratum
- for analyses by weight cohort, the CMH risk difference and the associated 95% CI, stratified by previous TNF-alpha antagonist therapy status

For binary endpoints, any subject with missing information for determination of endpoint status will be considered as a nonremitter/nonresponder at the respective timepoint. Subjects who had to switch to a higher dose due to loss of clinical response or disease worsening and subjects who received rescue corticosteroids will be considered nonremitters/nonresponders.

6.1.4 Definition of Study Days

Day 1 will be defined as the day of first study drug administration from the induction period, as recorded on the electronic case report form (eCRF) dosing page.

Study day will be calculated relative to the date of the first dose of study drug. Study days prior to the first dose of study drug will be calculated as:

- Date of assessment/event – Date of first dose of study drug.

Study days on or after the first dose of study drug will be calculated as:

- Date of assessment/event – Date of first dose of study drug + 1.

6.1.5 Analysis Approach for Time-to-Event Variables

Not applicable.

6.1.6 Calculations of Efficacy Endpoints

6.1.6.1 Conventions for Calculation of PUCAI

The PUCAI is composed of 6 clinical items (see Section 9.5 for details) for a daily average of last two days (Turner, 2007).

To calculate the PUCAI total score for a study visit, sum the 6 subscores at that study visit. If any of the 6 subscores is missing, the PUCAI total score cannot be calculated and the PUCAI total score for that study visit will be set to missing.

The PUCAI score ranges from 0 to 85; UC disease activity is categorized based on the PUCAI as follows:

- Remission: total score less than 10 points.
- Mild disease activity: total score between 10 and 30 points, inclusive.
- Moderate disease activity: total score between 35 and 60 points, inclusive.
- Severe disease activity: total score of 65 points or greater.

Clinical response is defined as a PUCAI change from baseline of ≥20 points.

6.1.6.2 Conventions for Calculation of Mayo Score

The Mayo scoring system is a composite index of 4 disease activity variables (see Section 9.4 for details):

- Stool frequency
- Rectal bleeding
- Findings on endoscopy
- Physician's global assessment (PGA)

Each variable is scored individually on an integer scale of 0 to 3, with higher scores indicating greater disease activity.

In addition to the complete Mayo score, two further variants of the Mayo score will be derived. The partial Mayo score is calculated analogously but excludes the endoscopy subscore. The modified Mayo score is calculated analogously but excludes the PGA subscore. The table below displays the 4 variables will be used for the various Mayo score calculations.

Table 6.a Components to the different variants of the Mayo Score

Score	Stool frequency	Rectal bleeding	Findings on endoscopy[a]	Physician's global assessment (PGA)
Modified Mayo	X	X	X	
Partial Mayo	X	X		X
Complete Mayo	X	X	X	X

[a] Endoscopy collected at Screening, Week 14 and Week 54 only.

All subscores should be rounded to the nearest integers prior to calculation of total score. The patient diary entries (i.e. stool frequency and rectal bleeding) on the day prior to, the day of and the day after endoscopy must not be used towards subscore calculations because of the required bowel preparation for the procedure.

The Mayo score and its variants described above will be derived using the different algorithms (approaches) for calculation of the clinical symptom subscores for stool frequency and rectal bleeding based on the patient reported diary data.

- Approach 1: Draft FDA UC Guidance 2016 approach
  (primary approach for analysis of Mayo-score-based efficacy endpoints)
- Approach 2: Gemini approach (Sensitivity 1)
- Approach 3: Draft FDA UC Guidance 2022 approach (Sensitivity 2)

Details for these 3 algorithms are described in the following subsections.

6.1.6.3 Gemini Approach

The complete Mayo score, partial Mayo score, and modified Mayo score for each subject will be calculated using the conventions applied for the calculation of Mayo score in the GEMINI UC study (C13006).

1. Use the date of the visit where PGA was performed to identify analysis visit using the analysis visit windowing rules defined in Section 9.2.4.

Identify PGA result (subscore).

Identify the endoscopy subscore (based on adjudicated score) using the analysis visit windows.

Calculate rectal bleeding subscore and stool frequency subscore:

a) Select the diary data completed by the patient from 7 days prior to the visit date identified in (1).

b) Merge in endoscopy dates (including dates of attempted endoscopy) and set diary data one day prior, on the day and one day after the endoscopy to missing.

c) For Baseline, if less than 3 days of data remain then a subscore cannot be calculated. Otherwise, sum the 3 most recent non-missing results and divide by 3. Patients who have less than 3 days of diary data at Baseline are not eligible for enrollment and the subscore will be considered missing.

d) For post-Baseline visits, sum the 3 most recent non-missing results and divide by 3. If only 2 non-missing results remain, then sum the 2 most recent non-missing results and divide by 2. If less than 2 days of diary data are available, the patient will be categorized as a non-responder and the subscore will be considered missing.

Calculate total scores:

a) For complete Mayo score, sum the PGA subscore, endoscopy subscore, rectal bleeding subscore and stool frequency subscore. All 4 subscores must be available.

b) For partial Mayo score, sum the PGA subscore, rectal bleeding subscore and stool frequency subscore. All 3 subscores must be available.

c) For modified Mayo score, sum the endoscopy subscore, rectal bleeding subscore and stool frequency subscore. All 3 subscores must be available.

Table 6.b provides examples of calculated Mayo scores using various Diary scenarios (Excluding Baseline):

Table 6.b Examples of Diary Subscore Entries and Corresponding Subscore Derivation per GEMINI Approach

Example	Diary Day[a]							Valid Days for Calculation of Subscore	Average Subscore	Final Subscore
	-7	-6	-5	-4	-3	-2	-1			
Diary #1	X	S[b]	X	2	**3**	**0**	**1**	-1, -2, -3	1.33	1
Diary #2	**3**	X	S[b]	X	**1**	M[c]	**2**	-1, -3, -7	2	2
Diary #3	S[b]	X	**3**	M[c]	M[c]	M[c]	**0**	-1, -5	1.5	2
Diary #4	4	**4**	X	S[b]	X	**3**	**3**	-1, -2, -6	3.33	3
Diary #5	2	**3**	**4**	**4**	X	S[b]	X	-4, -5, -6	3.67	4
Diary #6	**2**	M[c]	M[c]	X	S[b]	X	**2**	-1, -7	2	2
Diary #7	M[c]	3	X	S[b]	X	M[c]	M[c]	Missing	N/A	Missing

[a] Days are named relative to Day 1, which is the Day of the Study Visit; [b] Endoscopy; [c] Missing

6.1.6.4 Draft FDA UC Guidance 2016 Approach

The complete Mayo, modified Mayo, and partial Mayo scores and the corresponding efficacy endpoints will also be derived according to the draft FDA UC Guidance (August 2016). This FDA guidance approach will serve as the main method to calculate Mayo scores and to derive all Mayo score-based efficacy endpoints.

1. Use the date of the visit where PGA was performed to identify analysis visit using the analysis visit windowing rules from section 9.2.4.
2. Identify PGA result (subscore).
3. Identify the endoscopy subscore (based on adjudicated data) using the visit windows.
4. Calculate rectal bleeding subscore and stool frequency subscore:

a) Select the diary data completed by the patient from 7 days prior to the visit date identified in (1).

b) Merge in endoscopy dates (including dates of attempted endoscopy) and set diary data one day prior, on the day and one day after the endoscopy to missing.

c) For Baseline, if less than 3 days of data remain then a subscore cannot be calculated. Otherwise, sum the 3 most recent non-missing results and divide by 3. Patients who have less than 3 days of diary data at Baseline are not eligible for enrollment and the subscore will be considered missing.

d) For post-Baseline visits, sum the 3 most recent consecutive non-missing results and divide by 3. For patients who do not have 3 consecutive days of non-missing diary data but have at least 4 days of data available in the last 7-day period prior to the visit, the non-missing scores from the total count of available days in the last 7-day period will be averaged. If less than 3 consecutive days or 4 days of diary data in the last 7-day period are available, the patient will be categorized as a non-responder and the subscore will be considered missing.

Calculate total scores:

a) For complete Mayo score, sum the PGA subscore, endoscopy subscore, rectal bleeding subscore and stool frequency subscore. All 4 subscores must be available.

b) For partial Mayo score, sum the PGA subscore, rectal bleeding subscore and stool frequency subscore. All 3 subscores must be available.

c) For modified Mayo score, sum the endoscopy subscore, rectal bleeding subscore and stool frequency subscore. All 3 subscores must be available.

Table 6.c provides examples of calculated Mayo scores using various Diary scenarios (Excluding Baseline) per draft FDA UC guidance:

Table 6.c Examples of Diary Subscore Entries and Corresponding Subscore Derivation per draft FDA UC 2016 Guidance

Example	Diary Day[a]							Valid Days for Calculation of Subscore	Average Subscore	Final Subscore
	-7	-6	-5	-4	-3	-2	-1			
Diary #1	X	S[b]	X	2	**3**	**0**	**1**	-1, -2, -3	1.33	1
Diary #2	**3**	X	S[b]	X	1	M[c]	2	Missing	N/A	Missing
Diary #3	S[b]	X	**3**	M[c]	M[c]	M[c]	**0**	Missing	N/A	Missing
Diary #4	**4**	**4**	X	S[b]	X	**3**	**3**	-1, -2, -6, -7	3.5	4
Diary #5	2	**3**	**4**	**4**	X	S[b]	X	-4, -5, -6,	3.67	4
Diary #6	2	M[c]	M[c]	X	S[b]	X	2	Missing	N/A	Missing
Diary #7	M[c]	3	X	S[b]	X	M[c]	M[c]	Missing	N/A	Missing

[a] Days are named relative to Day 1, which is the Day of the Study Visit; [b] Endoscopy; [c] Missing

6.1.6.5 Draft FDA UC Guidance 2022 Approach

The complete Mayo, modified Mayo, and partial Mayo scores and the corresponding efficacy endpoints will also be derived according to the draft FDA UC Guidance (April 2002).

1. Use the date of the visit where PGA was performed to identify analysis visit using the analysis visit windowing rules from section 9.2.4.
2. Identify PGA result (subscore).
3. Identify the endoscopy subscore (based on adjudicated data) using the visit windows.
4. Calculate rectal bleeding subscore and stool frequency subscore:

d) Select the diary data completed by the patient from 7 days prior to the visit date identified in (1).

e) Merge in endoscopy dates (including dates of attempted endoscopy) and set diary data one day prior, and on the day of the endoscopy to missing.

f) For Baseline and post-Baseline visits:

- For patients who have at least 3 consecutive days of non-missing diary data available in the last 7-day period prior to the visit, the non-missing scores from the total count of available days in the last 7-day period will be averaged.
- For patients who do not have 3 consecutive days of non-missing diary data but have at least 4 non-consecutive days of data available in the last 7-day period prior to the visit, the non-missing scores from the total count of available days in the last 7-day period will be averaged.
- If at least 3 consecutive days or 4 non-consecutive days of diary data in the last 7-day period are not available, the patient will be categorized as a non-responder and the sub score will be considered missing.

Calculate total scores:

g) For complete Mayo score, sum the PGA subscore, endoscopy subscore, rectal bleeding subscore and stool frequency subscore. All 4 subscores must be available.

h) For partial Mayo score, sum the PGA subscore, rectal bleeding subscore and stool frequency subscore. All 3 subscores must be available.

i) For modified Mayo score, sum the endoscopy subscore, rectal bleeding subscore and stool frequency subscore. All 3 subscores must be available.

Table 6.d provides examples of calculated Mayo scores using various Diary scenarios (Excluding Baseline) per draft FDA UC guidance:

Table 6.d Examples of Diary Subscore Entries and Corresponding Subscore Derivation per draft FDA UC 2022 Guidance

Example	Diary Day[a]							Valid Days for Calculation of Subscore	Average Subscore	Final Subscore
	-7	-6	-5	-4	-3	-2	-1			
Diary #1	X	S[b]	M[c]	2	**3**	**0**	**1**	-1, -2, -3, -4	1.50	2
Diary #2	**3**	X	S[b]	2	1	M[c]	2	-1, -3, -4, -7	2	2
Diary #3	S[b]	1	**3**	M[c]	M[c]	M[c]	**0**	Missing	N/A	Missing
Diary #4	**4**	**4**	X	S[b]	1	**3**	**3**	-1, -2, -3, -6, -7	3.0	3
Diary #5	2	**3**	**4**	**4**	X	S[b]	1	-1, -4, -5, -6, -7	2.8	3
Diary #6	2	M[c]	M[c]	X	S[b]	3	2	Missing	N/A	Missing
Diary #7	M[c]	3	X	S[b]	3	M[c]	M[c]	Missing	N/A	Missing

[a] Days are named relative to Day 1, which is the Day of the Study Visit; [b] Endoscopy; [c] Missing

6.1.6.6 Impact-III

The IMPACT-III questionnaire is a self-reported measure with 35 closed questions encompassing 6 domains: Bowel Symptoms (7 items), Systemic Symptoms (3 items), Social Functioning (12 items), Body Image (3 items), Treatment/Interventions (3 items), and Emotional Functioning (7 items). The IMPACT-III uses a 5-point Likert scale ranging from 1 to 5 for all answers. The outcome score ranges from 35 to 175, with higher scores suggesting better quality of life. IMPACT-III (where translations are available) will be administered to subjects aged 9 to 17 years at the time of first dose of study drug and at Weeks 14 and 54.

Refer to Appendix 9.6 for details.

6.2 Disposition of Subjects

The primary reason for screen failure will be summarized for all subjects that screen failed.

The number of patients receiving study treatment, who completed or prematurely withdrew from the study, and the reasons for any premature withdrawal, during induction, maintenance, and

induction and maintenance respectively, will be summarized by weight group, dose group, and overall. These summaries will be performed for the ITT and ITT-M analysis sets. The number of patients eligible for each analysis set will be summarized for the ITT and ITT-M as applicable.

Time on the study will be presented by weight group and dose group. Time on the study (in weeks) will be derived as (End of Study Date - Date of Informed Consent/Assent +1) / 7. Summaries will be performed on the ITT and ITT-M analysis set.

Significant protocol deviations will be summarized for the ITT and ITT-M analysis sets.

Enrollment will be summarized by weight group, region, country and site for the ITT set.

6.3 Demographic and Other Baseline Characteristics

6.3.1 Demographics

Demographic information to be obtained may include age, sex, Hispanic ethnicity, race as described by the subject or parent/guardian, and geographic region.

Demographic information will be summarized for the ITT and ITT-M analysis set.

6.3.2 Medical History and Concurrent Medical Conditions

Medical history includes all medical conditions with a stop date prior to the first dose. A concurrent medical condition is a condition which occurs on or after the date of first dose, including those started before but which are ongoing on the day of first dose.

The coding dictionary to be used for medical history and concurrent medical conditions is MedDRA. Prior medical history and concurrent medical conditions will be summarized by system organ class and preferred term by weight cohort and overall.

Analysis from this section will be done for ITT analysis set.

6.3.3 Baseline Disease Characteristics

Baseline disease characteristics will be summarized based on ITT and ITT-M analysis set. The summaries will include descriptive statistics of all variables in the table below.

Table 6.e Baseline Disease Characteristics

Baseline Characteristics	Summarized as	Categories
Disease duration (in years) =(Informed Consent date – Diagnosis by Physician +1) /365.25 [1]	Continuous and Categorical	< 1 year ≥ 1 to < 3 years ≥ 3 to < 7 years ≥ 7 years
Baseline fecal calprotectin (µg/g)	Continuous and Categorical	≤ 250 µg/g > 250 µg/g and ≤ 500 µg/g > 500 µg/g

Table 6.e Baseline Disease Characteristics

Baseline Characteristics	Summarized as	Categories
Baseline CRP (mg/L)	Continuous and Categorical	≤ 1 mg/L > 1 mg/L and ≤ 3 mg/L > 3 mg/L and ≤ 10 mg/L > 10 mg/L
Baseline CRP (mg/L) - Category 2	Categorical	<=3 mg/L, >3 mg/L and ≤ 5 mg/L >5 mg/L and ≤ 10 mg/L >10 mg/L
TNF-alpha antagonist therapy status	Categorical	Naïve, Exposed/Failed
Baseline complete Mayo score	Continuous	
Baseline disease activity based on complete Mayo	Categorical	Mild (<6), Moderate (6 to 8), Severe (9 to 12)
Baseline partial Mayo score	Continuous	
Baseline disease activity based on partial Mayo	Categorical	<2, Mild (2-4), Moderate (5-6), Severe (7-9)
Baseline modified Mayo score	Continuous	
Baseline PUCAI	Continuous	
Baseline disease activity based on PUCAI	Categorical	< 10 Mild (10-30, inclusive), Moderate (35-60, inclusive), Severe (65 or greater)
Baseline stool frequency subscore for Mayo	Categorical	0, 1, 2, 3
Baseline rectal bleeding subscore for Mayo	Categorical	0, 1, 2, 3
Baseline endoscopic subscore from Mayo	Categorical	0, 1, 2, 3
Baseline physician's global assessment subscore from Mayo	Categorical	0, 1, 2, 3
Any acute exacerbations associated within the past 12 months	Categorical	Yes/No
Number of acute exacerbations	Continuous	
Any hospitalizations for UC within the past 12 months	Categorical	Yes/No
Number of hospitalizations	Continuous	
Disease Localization [2]	Categorical	Proctosigmoiditis, Left Sided Colitis, Extensive Colitis, Pancolitis
Any surgery for Ulcerative Colitis	Categorical	Yes/No
Number of surgeries	Continuous	
Extraintestinal Manifestations		
Arthritis/Arthralgia	Categorical	Yes/No
Iritis/Uveitis	Categorical	Yes/No

Table 6.e Baseline Disease Characteristics

Baseline Characteristics	Summarized as	Categories
Erythema nodosum	Categorical	Yes/No
Pyoderma gangrenosum	Categorical	Yes/No
Aphthous stomatitis	Categorical	Yes/No
Abscess	Categorical	Yes/No
Fever over 37.8 degrees Celsius during the past week	Categorical	Yes/No
Other Extraintestinal Manifestation	Categorical	Yes/No
Corticosteroid use over the last 12 months	Categorical	Yes/No
Estimated number of weeks of corticosteroid use over the last 12 months	Continuous	
Baseline Corticosteroid Use [3]	Categorical	Yes/No

Mayo scores in this table will be done by FDA Method only.

[1] The missing dates for diagnosis will be imputed with June for month and 15th for date.

[2] Disease Localization will be summarized based on the maximum extent of localization (Proctosigmoiditis < Left Sided Colitis < Extensive Colitis < Pancolitis).

[3] Baseline Corticosteroid Use is based on steroids that was used prior to first dose and ongoing at Day 1.

6.4 Medication History, Concomitant Medications and Procedures

Prior biologic history for the treatment of UC is defined as prior biologic medications stopped at or prior to signing of informed consent.

Prior medication given between 35 days prior to ICF date and Day 1 (inclusive of medication stopped on Day 1) will be summarized by weight group and overall.

Medication history is defined as any medication relevant to eligibility criteria stopped at or within 1 month prior to signing of informed consent.

Concomitant medication is defined as all medication given on or after Day 1 up to end of study.

Prior biologic history, medication history, prior medications and concomitant medications will be summarized by WHO Drug therapeutic class and preferred medication name, by weight cohort and overall. The number and percentage of patients will be provided for these analyses.

Analysis from this section will be summarized for ITT analysis set.

The concomitant medications will be also summarized for ITT-M.

Concomitant procedures, defined as any procedure that happened on or after Day 1 to end of study, will be also summarized by MedDRA Terms system organ class (SOC) and preferred term (PT) for ITT analysis set and ITT-M analysis set.

6.5 Efficacy Analysis

No formal hypothesis testing will be performed.

The efficacy analysis for the maintenance period will be performed by dose group and weight cohort; weight cohorts analyzed are high weight cohort (HWC, ≥ 30kg), low weight cohort (LWC, including weight groups 10 – 15 kg and > 15 to < 30 kg), and overall (LWC+HWC), unless otherwise specified.

6.5.1 Primary Endpoint Analysis

6.5.1.1 Derivation of Endpoint

Using the draft FDA UC Guidance (2016) approach, the primary endpoint is clinical remission at Week 54, where clinical remission based on the modified Mayo score (mMS-clinical remission) is defined as the sum of the following:

- stool frequency subscore 0 to 1 and a decrease of 1 or more from baseline; and
- rectal bleeding subscore of 0; and
- endoscopy subscore 0 to 1 (modified so that a score of 1 does not include friability).

6.5.1.2 Main Analytical Approach

The primary endpoint, mMS-clinical remission at Week 54, will be analyzed using the analysis approach described in section 6.1.3 using ITT-M analysis set in accordance with the draft FDA UC Guidance (2016) approach (section 6.1.6.4) and using the stratification factors from the IRT with consideration of ICE as defined in 1.3.

6.5.1.3 Sensitivity Analysis

Sensitivity analyses for clinical remission at Week 54 will be done using mITT and PPS-M analysis sets with consideration of ICE as defined in 1.3.

If there is any discrepancy between the IRT and EDC stratification, another sensitivity analysis will be done for the primary endpoint with consideration of ICE as defined in 1.3, with the ITT-M analysis set using the stratification factors from the EDC.

Furthermore, sensitivity analyses for the primary endpoints will be done based on the ITT-M analysis set using the following two alternative algorithms to derive the clinical symptoms (patient-reported) Mayo subscore to determine the modified Mayo score:

5. GEMINI method (section 6.1.6.3) with consideration of ICE as defined in 1.3.
6. Algorithm per the FDA UC Guidance (2022) (section 6.1.6.5) with consideration of ICE as defined in 1.3.

6.5.1.4 Supplemental Analyses

Supplemental analysis for clinical remission at Week 54 will be done using the ITT analysis set with consideration of ICE as defined in 1.3. This supplemental analysis will be done only based the Mayo score derived based on the FDA UC Guidance (2016) approach. Subjects in the ITT analysis set, who were not randomized into the maintenance phase, will be considered non-responders.

6.5.2 Secondary Endpoints Analysis

6.5.2.1 Secondary Endpoints

Table 6.f Secondary Efficacy Endpoints

Endpoint	Secondary	Analysis Set	Timepoint
Secondary Endpoints for Week 14	Clinical remission based on modified Mayo score [mMS-clinical remission]	ITT	Week 14
	Endoscopic response at Week 14 [Endo response]		Week 14
Secondary Endpoints for Week 54	Sustained clinical remission based on modified Mayo score at both Week 14 and Week 54 [Sustained mMayo clinical remission]	ITT-M [1] ITT	Week 54
	Sustained endoscopic remission based on Week 14 and Week 54 [Sustained endo remission]		Week 54
	Endoscopic response at Week 54 [Endo response]		Week 54
	Clinical remission based on complete Mayo score [cMS clinical remission]		Week 54
	Sustained clinical response based on complete Mayo score at both Week 14 and Week 54 [cMS Sustained clinical response]		Week 54
	Corticosteroid-free clinical remission [CS-free mMS-clinical remission] [2]	ITT-M [1]	Week 54
Other Secondary Endpoints	Clinical response based on partial Mayo score [pMS Clinical response]	ITT	Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54
	Clinical remission based on partial Mayo score [pMS Clinical remission]	ITT	Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54
	Change from baseline in weight and linear growth	ITT-M	Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54
	Change from baseline in weight and linear growth z-score	ITT-M	Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54
	Tanner Stage V	ITT-M	Week 54

[1] The 95% CI based on CMH approach will be included, as discussed in 6.1.3.

[2] Corticosteroid-free endpoints will be analyzed for subjects with baseline corticosteroid use.

6.5.2.2 Derivation of Endpoints

- Clinical remission at Week 14, where clinical remission is as defined in Section 6.5.1.1.
- Sustained clinical remission at Week 54, where a subject achieves sustained clinical remission if he or she achieved clinical remission (as defined by primary endpoint) at Week 14 and at Week 54 with consideration of ICE as defined in 1.3.
- Sustained endoscopic remission, defined as Mayo endoscopic score (MES) of ≤1 point, at Week 14 and at Week 54 with consideration of ICE as defined in 1.3.
- Endoscopic response, defined as a decrease in MES ≥1 point at Week 14.
- Endoscopic response, defined as a decrease in MES ≥1 point at Week 54 with consideration of ICE as defined in 1.3.
- Corticosteroid-free clinical remission at Week 54, where a subject achieves corticosteroid-free clinical remission at Week 54 if he or she meets the definition described in the primary endpoint and was off corticosteroids at least 12 weeks prior to Week 54 with consideration of ICE as defined in 1.3.
    - The corticosteroid-free endpoint will be done amongst subjects with baseline corticosteroid use.
- Clinical remission based on complete Mayo score at Week 54, where a subject achieves clinical remission if he or she achieved a complete Mayo score ≤2 points with no individual subscore >1 at Week 54 with consideration of ICE as defined in 1.3.
- Sustained clinical response of subjects at Weeks 14 and 54, where a subject meets clinical response if he or she has a reduction in complete Mayo score (see Section 9.4) of ≥3 points and ≥30% from baseline with an accompanying decrease in rectal bleeding subscore of ≥1 point or absolute rectal bleeding subscore of ≤1 point with consideration of ICE as defined in 1.3.
- Clinical response at Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54 where a subject achieves clinical response if he or she meets the following definition:

Maintenance of a reduction of ≥2 points and ≥25% from the baseline partial Mayo score, including a ≥1-point decrease in the Mayo stool frequency subscore and a ≥1-point reduction in the rectal bleeding subscore or absolute rectal bleeding subscore of ≤1 point

- Shift in Tanner stage at Week 54 compared with baseline, each domain separately.
- Change in Baseline weight calculated as:

Weight at Week XX – Weight at Baseline

- Change from baseline in weight and linear growth z-score

6.5.2.3 Main Analytical Approach

All binary endpoints listed above with derivation defined in section 6.5.2.2 will be analyzed using the analysis approach described in section 6.1.2 and 6.1.3. For endpoints based on the Mayo score, the Mayo scores will be derived according to the draft FDA UC Guidance (2016) as described in 6.1.6.4.

The change from baseline in weight and linear growth z-scores will be presented with 95% CIs.

The shift table in Tanner stage at Week 54 compared with baseline will be presented for each of the domains, separately.

The analysis set for each analysis for efficacy endpoints are defined in section 6.5.2 and for immunogenicity endpoints are defined in section 6.9.1.

6.5.2.4 Sensitivity Analysis

Sensitivity analyses will be done using the following two alternative algorithms to derive the clinical symptoms (patient-reported) Mayo subscore:

1. GEMINI method (section 6.1.6.3)
2. Algorithm per the FDA UC Guidance (2022) (section 6.1.6.5)

These sensitivity analyses will be conducted for the following secondary endpoints:

- clinical remission based on modified Mayo score at Week 14 (ITT)
- clinical response based on modified Mayo score at Week 14 (ITT)
- sustained clinical remission based on modified Mayo score at Week 14 and Week 54 (ITT-M) with consideration of ICE as defined in 1.3.
- sustained endoscopic remission at Week 14 and Week 54 (ITT-M) with consideration of ICE as defined in 1.3.
- endoscopic response at Week 54 (ITT-M) with consideration of ICE as defined in 1.3.
- corticosteroid-free clinical remission at Week 54 (ITT-M) with consideration of ICE as defined in 1.3.
- clinical remission based on complete Mayo score at Week 54 (ITT-M) with consideration of ICE as defined in 1.3.
- sustained clinical response based on complete Mayo score at Week 14 and Week 54 (ITT-M) with consideration of ICE as defined in 1.3.

6.5.2.5 Supplementary Analyses

Supplemental analyses will be done for the following binary endpoints at Week 54 using the ITT analysis set with consideration of ICE as defined in 1.3. For the Mayo-score-based endpoints, the supplemental analyses will be done only for the Mayo scores based on FDA UC Guidance

(2016) approach. Subjects in the ITT analysis set, who weren't randomized into maintenance phase, will be considered non-responders.

- Sustained clinical remission based on modified Mayo score
- Sustained endoscopic remission
- Endoscopic response
- Clinical remission based on complete Mayo score
- Sustained clinical response based on complete Mayo score

6.5.3 Exploratory Endpoints Analysis

The following exploratory endpoints will be summarized descriptively using the analysis approach described in section 6.1.2. Mayo score calculation for Exploratory Analysis will be done based on FDA UC Guidance (2016) approach only.

- Change from baseline in CRP, albumin, and fecal calprotectin based on the ITT, by clinical response at Week 14, by visit for Weeks 2, 6, 10, and 14 by weight group at Day 1.
- Change from baseline in CRP, albumin, and fecal calprotectin based on the ITT-M by visit for Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54 by weight group and dose group at Week 14.
- Change from baseline in IMPACT-III total and subscale scores at Weeks 14 and 54 for subjects aged 9 to 17 years at the time of the first dose for Weeks 14 and Week 54 based on the ITT-M by weight group and dose group at Week 14. The corresponding 95% CIs will be presented.
- Change from baseline in partial Mayo (calculated using the FDA 2016 algorithm) and PUCAI scores and subscores based on the ITT by visit for Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54 by weight group at Day 1.

The following binary exploratory endpoints will be summarized by weight group and dose group at Week 14 using the analysis approach described in section 6.1.3.

- Clinical remission based on PUCAI score at Week 54, defined as a PUCAI score <10 at Week 54 based on ITT-M.
- Clinical response based on PUCAI, defined as a ≥20-point decrease from baseline in the PUCAI score, for Weeks 2, 6, 10, 14, 22, 30, 38, 46, and 54 based on the ITT-M.

6.5.4 Subgroup Analyses

Subgroup analyses will be performed for the two efficacy endpoints, with consideration of ICE as defined in 1.3, clinical remission at Week 54 (based on modified Mayo score) and sustained clinical remission at Week 54 for the ITT-M based on the following subgroups:

- TNF alpha antagonist therapy status (Naive, Exposure/Failure)
- Weight group (10 to 15 kg, >15 to < 30 kg, ≥ 30 kg)

- Age (≤ 12 years, > 12 years)
- Gender (male, female)
- Race
- Region
- Disease duration (< 1yr, 1 to <3 yrs, ≥ 3 yrs)
- Disease localization (proctosigmoiditis, left sided, extensive, pancolitis)
- Ethnicity (Hispanic or Latino vs not Hispanic or Latino).

6.5.5 Other Parameters for Analyses

The following parameters will be summarized descriptively by weight group and dose group using the analysis approach described in Section 6.1.3 based on the ITT. Mayo score calculation for Other Parameters for Analysis will be done based on FDA UC Guidance (2016) approach only.

Parameters	Derived by	Analysis Set	Timepoint
Clinical response based on modified Mayo score [mMs-clinical response]	Defined as a reduction of ≥2 points and ≥30% from the baseline modified Mayo score with an accompanying decrease in rectal bleeding subscore of ≥1 point or absolute rectal bleeding subscore of ≤1 point	ITT	Week 14

6.6 Safety Analysis

The safety analysis set (SAF) and safety analysis set – maintenance (SAF-M) will be used for summaries in this section. No statistical inference will be made.

6.6.1 Adverse Events

Treatment-emergent AEs (TEAEs) are undesirable events not present prior to medical treatment or an already-present event that worsens either in intensity or frequency following the treatment, occurring from the first dose of study drug to the day of last dose of study drug + 126 days (accounting for defined as 5 times the half-life of vedolizumab).

6.6.1.1 Four Types of Adverse Event Analysis

There will be four sets of tables for the adverse events based on onset of TEAE and analysis set flag:

- All TEAEs with an onset during induction and maintenance phase summarized for the SAF
- All TEAEs with an onset during the induction period summarized for the SAF

- All TEAEs with an onset during induction and maintenance phase summarized for the SAF-M
- All TEAEs with an onset during maintenance phase (i.e. TEAEs after first dose of maintenance dose) summarized for the SAF-M

6.6.1.2 Headers for the Safety Analysis Based on the Analysis Set

The tables using the SAF set will be summarized by weight groups at Day 1 and its corresponding dose as shown below:

1) 10-15 kg – 150 mg
2) >15 to < 30 kg – 200 mg
3) LWC (10-15 kg – 150 mg, >15 to < 30 kg – 200 mg combined)
4) HWC (≥ 30 kg – 300 mg)
5) Overall

The tables using SAF-M set will be summarized by weight groups at Week 14 and the actual dose group (Low/High) as shown below. The subjects who have dose escalated during the maintenance period will be summarized in the actual first dose received. The SAF-M analysis will have the following subheadings and columns in parenthesis.

1) All weights combined (dose groups: Low/High/Overall)
2) Weight group 10- 15 kg (dose groups: Low/High/Overall)
3) Weight group: > 15 - < 30 kg (dose groups: Low/High/Overall)
4) Weight group: ≥ 30 kg (dose groups: Low/High/Overall)

6.6.1.3 Overview of TEAEs

For the overview of TEAEs, the following will be summarized for the count and proportions of subjects.

- TEAE
- Related TEAE
- Not Related TEAE
- Mild TEAEs
- Moderate TEAEs
- Severe TEAEs
- TEAE Leading to Discontinuation
- Serious TEAE
- Related Serious TEAE

- Not Related Serious TEAE
- Serious TEAE Leading to Discontinuation and Death

6.6.1.4 Adverse Event Analysis

The count and proportion of subjects with the following categories of TEAEs will be summarized by MedDRA system organ class (SOC), high level term (HLT) and preferred term (PT) will be summarized for the content below:

- TEAEs
- Drug-related TEAEs
- TEAEs leading to study drug discontinuation
- Serious TEAEs
- Drug-related serious TEAEs

These analyses will be done for each 4 types of the analysis outlined in section 6.6.1.1.

For the analysis that includes all AEs with an onset during induction and maintenance phase summarized by SAF, the exposure-adjusted incidence rates will be displayed for the content below:

- TEAEs
- Serious TEAEs
- Drug-related TEAEs

The exposure-adjusted incidence rate will be defined as the number of subjects with events per person-time, which will be presented per 100 Person-Years. The person-time is defined as the duration of exposure and follow up period in this study.

Person-time for Simplified Exposure-Adjusted Incidence Rate

For subjects who enrolled into the treatment cohort of the extension study:

[1] Time at risk (years) = (MAX (date of Week 54, date of last study assessment) – date of first dose + 1) / 365.25

For subjects who do not enroll into the treatment cohort of the extension study:

[2] Time at risk (years) = (MIN (date of last dose of study drug in parent study + 126 days, date of last contact in parent study) – date of first dose + 1) / 365.25

Person-time for Risk-Adjusted Incidence Rate

[3] Time at risk (years) = (date of first occurrence of event – first dose date + 1) / 365.25

For subjects with no event, the exposure time at risk would be calculated as formulas from [1] or [2].

6.6.1.5 Other Adverse Event Analysis

The count and proportion of subjects will be summarized by MedDRA system organ class, high level term and preferred term will be summarized for the content below.

For TEAEs during induction and maintenance, the below will be also analyzed.

- Pretreatment event (PTE), which is defined as any untoward medical occurrence in a clinical investigation subject who has signed informed consent to participate in a study but prior to administration of any study drug; it does not necessarily have to have a causal relationship with study participation in the SAF set.
- TEAEs among the low dose group subjects by onset of TEAEs compared to dose escalation in the SAF-M set.
- Treatment-emergent serious adverse event (TESAE) among the low dose group subjects by onset of TEAEs compared to dose escalation in the SAF-M set.
- TEAE by severity in the SAF set.
- TEAE of special interest (AESI) for each category of malignancies, serious infections, infusion-related reactions, hypersensitivity reactions, liver injury, and PML. See Appendix 9.7 for details in the SAF set.

For TEAEs during maintenance, the below will be also analyzed.

- TEAE of special interest (AESI) for each category of malignancies, serious infections, infusion-related reactions, hypersensitivity reactions, liver injury, and PML. See Appendix 9.7 for details in the SAF-M set.

The count and proportion will be summarized by preferred term for the content below:

- Most Frequent TEAEs (during induction and maintenance) occurring >= 3% of the subjects in the SAF set
- Most Frequent TEAEs (during induction and maintenance) occurring >= 3% of the subjects in the SAF-M set

6.6.1.6 Listings for Adverse Event

The following will be presented in Listings format.

- Adverse Events
- Related Adverse Events
- Adverse Events leading to discontinuation
- Serious Adverse Events
- Adverse Events resulting in death

6.6.2 Vital Signs

Change from baseline will also be summarized based on SAF for vital signs including weight, height, body temperature, respiratory rate, blood pressure, and pulse (beats per minute) following the analysis approach described in Section 6.1.2; CIs will not be provided.

Markedly Abnormal Vital Signs will be summarized based on 9.9 Criteria for Markedly Abnormal Values for Vital Signs.

6.6.3 Clinical Laboratory Results

The change from baseline will be presented for the following clinical laboratory results (SI) at each time point for SAF as described in Section 6.1.2 (CIs will not be provided). The change from baseline will be also presented for Conventional Units (CV) for hematology and clinical chemistry.

Markedly Abnormal Laboratory will be summarized for post-baseline data for SAF based on International System of Units (SI) units based on 9.8 Criteria for Identification of Markedly Abnormal Laboratory Values.

Table 6.g Clinical Laboratory

Hematology	Clinical Chemistry
RBC	ALT
WBC with differential	Albumin
Hemoglobin	Alkaline phosphatase
Hematocrit	AST
Platelets	Direct bilirubin
INR (if required)	Total bilirubin
	Total protein
	Creatinine
	Blood urea nitrogen
	Creatine kinase
	GGT
	Potassium
	Sodium
	Glucose
	Chloride
	Bicarbonate
	Calcium
	Amylase
	CRP
Other:	
Serum	Urine

Table 6.g Clinical Laboratory

Hematology	Clinical Chemistry
HIV test Hepatitis panel, including HBsAg, HBcAb, HBsAb, and anti-HCV QuantiFERON test [a] AVA Female subject only: beta hCG[b]	Female subjects only: hCG (urine pregnancy test) (Female subjects only who are menstruating or aged ≥11 years, whichever is younger) [b]
*Female subjects only who are menstruating (regardless of age) or aged ≥11 years [b]	
Stool:	
C difficile testing and toxin A and B Fecal calprotectin Ova and parasite evaluation Stool culture	

ALT: alanine aminotransferase; AST: aspartate aminotransferase; AVA: antivedolizumab antibody; CRP: C-reactive protein; GGT: γ-glutamyl transferase; hCG: human chorionic gonadotropin; INR: international normalized ratio; RBC: red blood cell; WBC: white blood cell.

[a] TB skin test can be performed as alternative for screening.

[b] Serum pregnancy analysis is performed at the screening visit and, Week 54/ET, and safety follow-up visits. Urine pregnancy analysis is performed at all other scheduled visits as specified in Schedule of Study Procedures

6.6.4 Other Safety Analysis

Abnormal Hepatic Liver Functions based on SI will be summarized as the following:

1. ALT will be summarized for: >3x ULN, >5x ULN, >10x ULN, and >3x ULN with Total Bilirubin >2x ULN.
2. AST will be summarized for >3x ULN, >5x ULN, >10x ULN, and >3x ULN with Total Bilirubin >2x ULN.
3. ALT or AST will be summarized for >3x ULN, >5x ULN, >10x ULN, and >3x ULN with Total Bilirubin >2x ULN.
4. ALT and AST will be summarized for >3x ULN, >5x ULN, >10x ULN, and >3x ULN with Total Bilirubin >2x ULN.
5. Alkaline phosphate will be summarized for >3x ULN, >3x ULN with ALT >3x ULN, and >3x ULN with AST >3x ULN,

6.6.5 Extent of Exposure and Compliance

The safety analysis set (SAF) will be used for all summaries in this section and will be summarized by weight group at induction.

The count and percentage of subjects with total number of completed infusions in induction phase. Completed infusions are defined as infusions where the total amount of study drug is infused.

Exposure during the maintenance phase will be summarized by count and percentage by low dose or high dose, within the context of the weight of induction.

The count and percentage of subjects who dose escalated, among those who had at least 1 dose of maintenance drug, will be summarized.

The extent of exposure will be calculated by the duration between the first and last dose of study drug plus 18 weeks (126 days) in order to account for the known duration of detectable vedolizumab serum concentration after the last dose, i.e.

Duration = Date of last dose of vedolizumab – Date of first dose of vedolizumab + 1 + 126 days.

Duration of exposure will be summarized within the context of the weight group at induction for SAF, SAF-M, and SAF-I, and will be reported for weeks and years of duration.

Compliance will be calculated as the proportion of actual infusions out of the total number of planned infusions and will be summarized by SAF, SAF-I, and SAF-M.

Compliance = [(actual infusions received) / total number of expected infusions based on study participation] * 100

If subject participated until Week 14, that subject's compliance would be based on 14 weeks of study participation.

Study drug administration data, duration of exposure, and compliance will be presented in data listings.

6.7 Pharmacokinetic, Pharmacodynamic, and Biomarker Analyses

6.7.1 Pharmacokinetic Analysis

Pharmacokinetic data will be analyzed by using PKAS, PKAS-M, and PKAS-I depending on the specific analysis below.

The PK parameters will be summarized by weight group and dose group using descriptive statistics (non-missing values, mean, SD, CV%, geometric mean, geometric mean CV%, median, first quartile, third quartile, minimum and maximum) as appropriate.

If a concentration value is below the lower limit of quantitation (<LLOQ):

- Individual concentration listings for the clinical study report (CSR) will list the value below LLOQ as BLQ (below the limit of quantification) and list the value above the ULOQ as ALQ (above limit of quantification).
- For summary statistics for trough serum concentrations such as mean, median, standard deviation (SD), coefficient of variation expressed in percent (CV%), minimum and maximum, concentration values above the ULOQ will be considered as missing and concentration values below the LLOQ will be imputed as zero.

- For summary statistics for trough serum concentrations such as geometric mean, geometric coefficient of variation expressed in percent (geoCV%), concentration values below the LLOQ and above the ULOQ will be treated as missing.

The measured serum concentrations of vedolizumab will be summarized by weight group and visit for PKAS-I for the induction phase, and for PKAS-M for the maintenance phase (low, high, low-high dose).

The mean serum concentration of vedolizumab over time will be plotted on linear and semi-logarithmic scale with a base of 10.

The measured serum concentrations of vedolizumab that are pre-dose are considered $C_{trough.}$ The C_{trough} at Week 14 is the pre-dose serum concentration at Week 14, and C_{trough} at Week 54 is the pre-dose serum concentration at Week 54.

The C_{trough} at Week 14 by clinical response (based on modified Mayo) and C_{trough} at Week 14 by endoscopic response will be summarized for PKAS analysis set.

The below clinical endpoint will be summarized by C_{trough} quartiles:

- clinical remission based on modified Mayo at Week 54

The C_{trough} will be summarized by the following efficacy endpoints:

- clinical remission based on modified Mayo at Week 54 (by weight group)
- clinical remission based on modified Mayo at Week 54 (by weight cohort)
- clinical response based on partial Mayo at Week 54 (by weight group)
- endoscopic response at Week 54 (by weight group)

Serum concentration data and PK parameters will also be listed.

For PKAS-M analysis, subjects who has dose escalated column will be counted from the visit that the subject has dose escalated.

Further analysis will be performed as deemed necessary and will not be reported in the clinical study report. Analysis details and results will be part of a separate standalone report.

6.7.2 Pharmacodynamic Analysis

Not applicable.

6.7.3 Biomarker Analysis

The Pearson correlation with trough PK levels of vedolizumab against CRP, serum albumin, fecal calprotectin will be displayed with the figures in scatter plot format.

Within the serum quartiles of Ctrough at Week 14 and Week 54, median values of CRP, serum albumin and fecal calprotectin at the respective visit will be calculated.

6.8 Patient Reported Outcomes (PROs) and Health Care Utilization Endpoints Analysis

6.8.1 PRO Analysis

Change from baseline in IMPACT-III (where translations are available) total and subscale scores at Weeks 14 and 54 for eligible subjects in ITT-M will be summarized as described in Section 6.1.2.

The details of the sub score are defined in Appendix 9.6

6.8.2 Health Care Utilization Analysis

Not applicable.

6.9 Other Analyses

6.9.1 Immunogenicity Analysis

Immunogenicity of vedolizumab will be summarized using the SAF, SAF-M, and SAF-I, depending on the specific analysis below, including all evaluable samples collected during the study, i.e. from baseline/pre-dose to the last subject's last assessment (including the safety follow-up). Missing AVA data will not be imputed.

6.9.1.1 Derivation of Endpoints

Definitions at sample level:

- Negative AVA sample: a sample that was evaluated as negative in the AVA screening assay or a sample that was determined as potentially positive in the AVA screening assay but the result was not confirmed in the AVA confirmatory assay.
- Positive AVA sample: a sample that was evaluated as positive in both the AVA screening and AVA confirmatory assays.
- Positive neutralizing AVA sample: a sample that was evaluated as positive in the neutralizing AVA assay.

Definitions at subject level:

- A negative AVA subject: a subject who has negative AVA results at all time points during the study, from baseline/pre-dose to the subject's last assessment (including the safety follow-up).
- A positive AVA subject: a subject who has at least 1 confirmed positive AVA result during the study, from baseline/pre-dose to the subject's last assessment (including the safety follow-up) and is further categorized as:
    - Transiently positive: a subject with at least 1 confirmed positive AVA sample and no consecutive positive AVA samples

        - Baseline positive only: a subject with a confirmed positive AVA sample at baseline/pre-dose and negative AVA samples at all post-dose visits; this includes subjects with a positive sample at baseline and a) negative samples at all post-dose time points (with evaluable AVA sample) or b) no evaluable post-dose AVA sample available.
    - Persistently positive: a subject with confirmed positive AVA samples at 2 or more consecutive visits
- A neutralizing AVA (nAVA) positive subject: subject with any positive neutralizing AVA result during the study from baseline/pre-dose to the subject's last assessment (including the safety follow-up)

6.9.1.2 Main Analytical Approach

The AVA data will be displayed based on weight groups at Day 1 for SAF-I and maintenance data will be based on weight group and dose group at Week 14 for SAF-M, as applicable. Subjects who shifted weight group between induction and maintenance will be displayed in overall weight group only. The summaries for the SAF will be presented based on weight groups at Day 1.

For PKAS-M and SAF-M analysis, subjects who have dose escalated will be counted in the dose escalated column from the visit that the subject has dose escalated; prior to dose escalation, they will be counted in the dose group that the subject has received at Week 14.

The subjects' overall AVA (negative, positive, transiently positive and persistently positive, baseline positive only (if applicable) and nAVA status (any neutralizing AVA positive) during the study will be summarized as follows:

- Overall AVA status in the study (SAF)
- Overall status by clinical remission (based on the modified Mayo score) status at Week 54 (SAF-M)
- Overall status by clinical remission (based on the modified Mayo score) status at Week 14 (SAF-I and SAF-M)
- Overall status by presence of investigator-defined infusion related reactions (SAF-I and SAF-M)
- Overall status by presence of hypersensitivity reactions (SAF-I and SAF-M)

Furthermore, the AVA status (negative, positive) and nAVA positive samples by visit will be summarized as follows:

- AVA status by visit based on the SAF by weight group at Day 1.
- AVA status by visit and titer categories (for AVA positive only) based on the SAF-I for induction data, and SAF-M for induction and maintenance data.

- AVA status by visit and titer/dilution factor (for AVA positive) based on the SAF-I for induction data, and SAF-M for induction and maintenance data.

The confirmed positive samples will be reported in 5-fold serial dilution factors (10, 50, 250, 1250, 6250, 31250, etc.) and by AVA titer category; titer categories are based on dilution factors and defined as low (≤ 50), moderate (250-1250) and high (≥ 6250).

Vedolizumab serum concentrations (ug/ml) by visit will be summarized by AVA Status (negative/positive) at the corresponding visits based on the PKAS-I for induction data, and PKAS-M for induction and maintenance data.

6.9.2 Comparison with External Data

Comparisons of efficacy between pediatric and adult populations will not be conducted and reported in the study CSR. Therefore, details of those comparison analyses will not be provided in this SAP. Bayesian borrowing for partial extrapolation of adult data and meta-analysis for pooling placebo rates will be outside the scope of this SAP, with analysis details provided in separate SAPs.

6.10 Interim Analyses

6.10.1 Initial Interim Analyses

An interim analysis (IA) may be conducted when at least 90 subjects enrolled in the ≥30 kg cohort have completed the study through Week 54 or prematurely discontinued. The purpose of this IA would be to evaluate safety, efficacy, and PK of vedolizumab treatment through Week 54. This IA may support regulatory interactions and would be performed by an unblinded team not directly involved in study activities. A separate (blinded) study team would continue to support the ongoing study activities.

The scope of the IA is outlined below.

- All LWC and HWC subjects enrolled for induction and who have completed the study though Week 54 or discontinued by the data cutoff date will be included in the IA.
- Subject disposition will be summarized for the ITT and ITT-M (see section 6.2).
- Demographic and other baseline characteristics will be summarized for the ITT-M as described in section 6.3.1 and section 6.3.3. The prior medical history and concurrent medical conditions and details of prior biologic history for the ITT will be summarized as described in section 6.3.2.
- For the primary efficacy endpoint clinical remission at Week 54, the following analyses will be done (see also section 6.5.1).
    - the primary analysis according to the main analytical approach for the ITT-M as described in section 6.5.1.2
    - the sensitivity analysis based on GEMINI approach as described in section 6.5.1.3 for the ITT-M

  - the supplementary analysis as described in section 6.5.1.4 for the ITT
  - subgroup analyses by prior anti-TNF use and by weight group as described in section 6.5.4 for the ITT-M

- For the secondary Week 14 efficacy endpoints, clinical remission based on modified Mayo score, endoscopic response, clinical response based on modified Mayo score, the following analyses will be done (see also section 6.5.2) for the ITT:
  - the analysis according to the main analytical approach as described in section 6.5.2.3
  - the sensitivity analysis based on GEMINI approach as described in section 6.5.2.4
- For the secondary Week 54 efficacy endpoints, sustained clinical remission based on modified Mayo score, sustained endoscopic remission, endoscopic response, corticosteroid-free clinical remission, clinical remission based on complete Mayo score, sustained clinical response based on complete Mayo score, the following analyses will be done (see also section 6.5.2) for the ITT-M:
  - the analysis according to the main analytical approach as described in section 6.5.2.3
  - the sensitivity analysis based on GEMINI approach as described in section 6.5.2.4
- The vedolizumab PK endpoints will be analyzed as described in section 6.7.1.
  - the measured serum concentration for the PKAS-I and PKAS-M
  - the serum trough concentration by clinical response at Week 14 for the PKAS set
  - the serum trough concentration by clinical remission based on mMS at Week 54 by weight group for the PKAS set
  - the serum trough concentration by clinical remission based on mMS at Week 54 by weight cohort for the PKAS set
  - the serum trough concentrations quartiles of vedolizumab by clinical remission based on mMS at Week 54 by weight group and dose group for the PKAS-M
- The AVA endpoints will be analyzed as described in section 6.9.1.
  - the overall AVA status will be summarized for the SAF
  - the AVA status by clinical remission at Week 54 for the SAF-M
  - the AVA status by clinical remission at Week 14 for the SAF-I and SAF-M
  - the serum concentration by AVA status for the SAF-I and SAF-M.
- All AEs with an onset during induction and maintenance phase summarized for the SAF as described in section 6.6.1
  - Overview of TEAEs as described in section 6.6.1.3.

- Summaries of events by MedDRA SOC, HLT and PT will be provided for below as described in section 6.6.1.4.
  - TEAEs
  - drug-related TEAEs
  - serious TEAEs
  - TEAEs leading to study drug discontinuations
  - drug-related serious TEAEs

- The exposure and compliance data will be summarized for the SAF as described in section 6.6.5.

6.10.2 Week 54 Analysis (Week 54 IA or Final Analysis)

The Week 54 analysis may be conducted when all enrolled subjects have completed the study through Week 54 or are prematurely discontinued. All efficacy and safety data collected up to Week 54/ET data cut would be summarized. This may support further regulatory interactions and/or submissions.

The Week 54 analysis will be considered a Week 54 IA if there are subjects remaining to complete the 18-week safety follow up at the Week 54 IA database lock. If there are no subjects pending to complete the 18-week safety follow up at the Week 54 final database lock, the Week 54 analysis will be considered the final analysis of the study.

Therefore, the Sponsor study team would be unblinded upon database lock (ie, Week 54 IA DBL or Week 54 final DBL) of this analysis. However, the study subjects, their legal representatives, and study physicians will be kept blinded until the last subject has completed the last study visit.

6.10.3 Final Analysis (aka CSR Addendum, if applicable)

If there are no subjects pending after the Week 54 database lock, the Week 54 analysis will be considered the final analysis and CSR addendum will not be needed.

If there are subjects remaining to complete the 18-week safety follow up after the Week 54 IA database lock, this additional data will be summarized as a CSR addendum as specified in section 6.10.3.1, as the final analysis.

This final analysis of the study will be conducted when all subjects completed their last study visit, and will be reported in the CSR addendum, if applicable.

6.10.3.1 Addendum Scope

The scope of this analysis will include the following summary tables for the SAF; all analyses will be done, which means pooled across dose groups:

- Subject disposition will be summarized for the ITT and ITT-M (see section 6.2).
- The concomitant medications and concomitant procedures will be summarized as described in 6.4.

- The below table data as described in section 6.7.1 will be summarized.
    - The measured serum concentrations of vedolizumab will be summarized by weight group and visit for PKAS-I for the induction phase, and for PKAS-M for the maintenance phase (low, high, low-high dose).
    - The mean serum concentration of vedolizumab over time will be plotted on linear and semi-logarithmic scale with a base of 10.
    - Serum concentration data and PK parameters will also be listed.
- The AVA endpoints will be analyzed as described in section 6.9.1.
- All AEs with an onset during induction and maintenance phase summarized for the SAF as described in section 6.6.1.
- All AEs with an onset during induction and maintenance phase summarized for the SAF-M as described in section 6.6.1.
- All further safety data including vital signs and laboratory results for the SAF as described in sections 6.6.2, 6.6.3, 6.6.4.
- The exposure and compliance data will be summarized for the SAF as described in section 6.6.5.

6.11 Data Monitoring Committee/Internal Review Committee/ Other Data Review Committees

There will be a Data Monitoring Committee (DMC). DMC Charter will specify details on the conduct and processes for the DMC.

6.12 Handling of Missing, Unused, and Spurious Data

6.12.1 Missing Date of Investigational Product

When the date of the last dose of study assigned treatment is missing for a subject in the safety set, all efforts should be made to obtain the date from the investigator. If it is still missing after all efforts, then the last visit date when study assigned treatment was returned will be used in the calculation of treatment duration.

6.12.2 Missing Date Information for Prior or Concomitant Medications (Therapies/Procedures)

For prior or concomitant medications, or procedure, incomplete (i.e., partially missing) start date and/or stop date will be imputed. When the start date and the stop date are both incomplete for a subject, impute the start date first.

6.12.3 Incomplete Start Date

The following rules will be applied to impute the missing numerical fields. If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

6.12.4 Missing Day and Month

- If the year of the incomplete start date is the same as the year of the date of the first dose of investigational product, then the day and month of the date of the first dose of investigational product will be assigned to the missing fields
- If the year of the incomplete start date is before the year of the date of the first dose of investigational product, then December 31 will be assigned to the missing fields
- If the year of the incomplete start date is after the year of the date of the first dose of investigational product, then 01 January will be assigned to the missing fields.

Missing Month Only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing Day Only

- If the month and year of the incomplete start date are the same as the month and year of the date of the first dose of investigational product, then the day of the date of the first dose of investigational product will be assigned to the missing day
- If either the year is before the year of the date of the first dose of investigational product or if both years are the same, but the month is before the month of the date of the first dose of investigational product, then the last day of the month will be assigned to the missing day
- If either the year is after the year of the date of the first dose of investigational product or if both years are the same, but the month is after the month of the date of the first dose of investigational product, then the first day of the month will be assigned to the missing day.

6.12.5 Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of investigational product is missing, then replace it with the last treatment visit date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

Missing Day and Month

- If the year of the incomplete stop date is the same as the year as of the date of the last dose of investigational product, then the day and month of the date of the last dose of investigational product will be assigned to the missing fields

- If the year of the incomplete stop date is before the year of the date of the last dose of investigational product, then 31 December will be assigned to the missing fields
- If the year of the incomplete stop date is after the year of the date of the last dose of investigational product, then 01 January will be assigned to the missing fields.

Missing Month Only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

Missing Day Only

- •If the month and year of the incomplete stop date are the same as the month and year of the date of the last dose of investigational product, then the day of the date of the last dose of investigational product will be assigned to the missing day
- If either the year is before the year of the date of the last dose of investigational product or if both years are the same, but the month is before the month of the date of the last dose of investigational product, then the last day of the month will be assigned to the missing day
- If either the year is after the year of the last dose of investigational product or if both years are the same, but the month is after the month of the date of the last dose of investigational product, then the first day of the month will be assigned to the missing day.

6.12.6 Missing Date Information for Adverse Events

For AEs with partial start dates, non-missing date parts will be used to determine if the AE is treatment-emergent or not. If a determination cannot be made using the non-missing date parts as to when the AE occurred relative to study drug administration, e.g. AE start year and month are the same as the year and month of the first dose of investigational product, then the AE will be classified as treatment-emergent.

To facilitate categorization of AEs as treatment emergent, imputation of dates can be used. For AEs, the default is to only impute incomplete (i.e., partially missing) start dates. Incomplete stop dates may also be imputed when calculation of the duration of an AE is required per the protocol. If imputation of an incomplete stop date is required, and both the start date and the stop date are incomplete for a subject, impute the start date first.

- AE with incomplete start date will be imputed following the same rule as described above for incomplete start date for the concomitant medications.
- AE with incomplete stop date will be imputed following the same rule as described above for incomplete stop date for the concomitant medications.

6.12.7 Missing Severity Assessment for Adverse Events

If severity is missing for an AE starting prior to the date of the first dose of investigational product, then a severity of “Mild” will be assigned. If the severity is missing for an AE starting on or after the date of the first dose of investigational product, then a severity of “Severe” will be

assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.

6.12.8 Missing Relationship to Investigational Product for Adverse Events

If the relationship to investigational product is missing for an AE starting on or after the date of the first dose of investigational product, a causality of "Related" will be assigned. The imputed values for relationship to double-blind investigational product will be used for incidence summaries.

7.0 REFERENCES

Shavit-Brunschwig Z, Ledder O, Focht G, Urlep D, Lev-Tzion R, Marcus D, et al. Vedolizumab is effective in real life pediatric inflammatory bowel disease: report from the prospective multi-centre VEDOKIDS cohort study. 14th Congress of ECCO- European Chrohn's and Colitis Organisation; March, 2019; Copenhagen.

Abitbul G, Marcus D, Focht G, Assa A, Levine A, Broide E, et al. Vedolizumab therapeutic drug monitoring may predict outcome in pediatric IBD (PIBD). J Pediatr Gastroenterol Nutr 2017;65(Suppl 1):S24. Abst EP39.

Dowd B, Jossen J, Dubinsky M. Vedolizumab effectiveness associated with deep remission in pediatric inflammatory bowel disease. 2017;65(Supplement 2):S148. North American Society for Pediatric Gastroenterology, Hepatology and Nutrition Annual Meeting, Abstract 322.

Singh N, Dubinsky M, Singh A, Check M, Rabizadeh S. Single center experience of longterm vedolizumab efficacy in pediatric patients with inflammatory bowel disease. J Pediatr Gastroenterol Nutr 2017;65(Suppl 2):S37.

Conrad MA, Stein RE, Maxwell EC, Albenberg L, Baldassano RN, Dawany N, et al. Vedolizumab therapy in severe pediatric inflammatory bowel disease. Inflamm Bowel Dis 2016;22(10):2425-31.

Singh N, Rabizadeh S, Jossen J, Pittman N, Check M, Hashemi G, et al. Multi-center experience of vedolizumab effectiveness in pediatric inflammatory bowel disease. Inflamm Bowel Dis 2016;22(9):2121-6.

Ledder O, Assa A, Levine A, Escher JC, de Ridder L, Ruemmele F, et al. Vedolizumab in paediatric inflammatory bowel disease: a retrospective multi-centre experience from the paediatric IBD porto group of ESPGHAN. J Crohns Colitis 2017;11(10):1230-7.

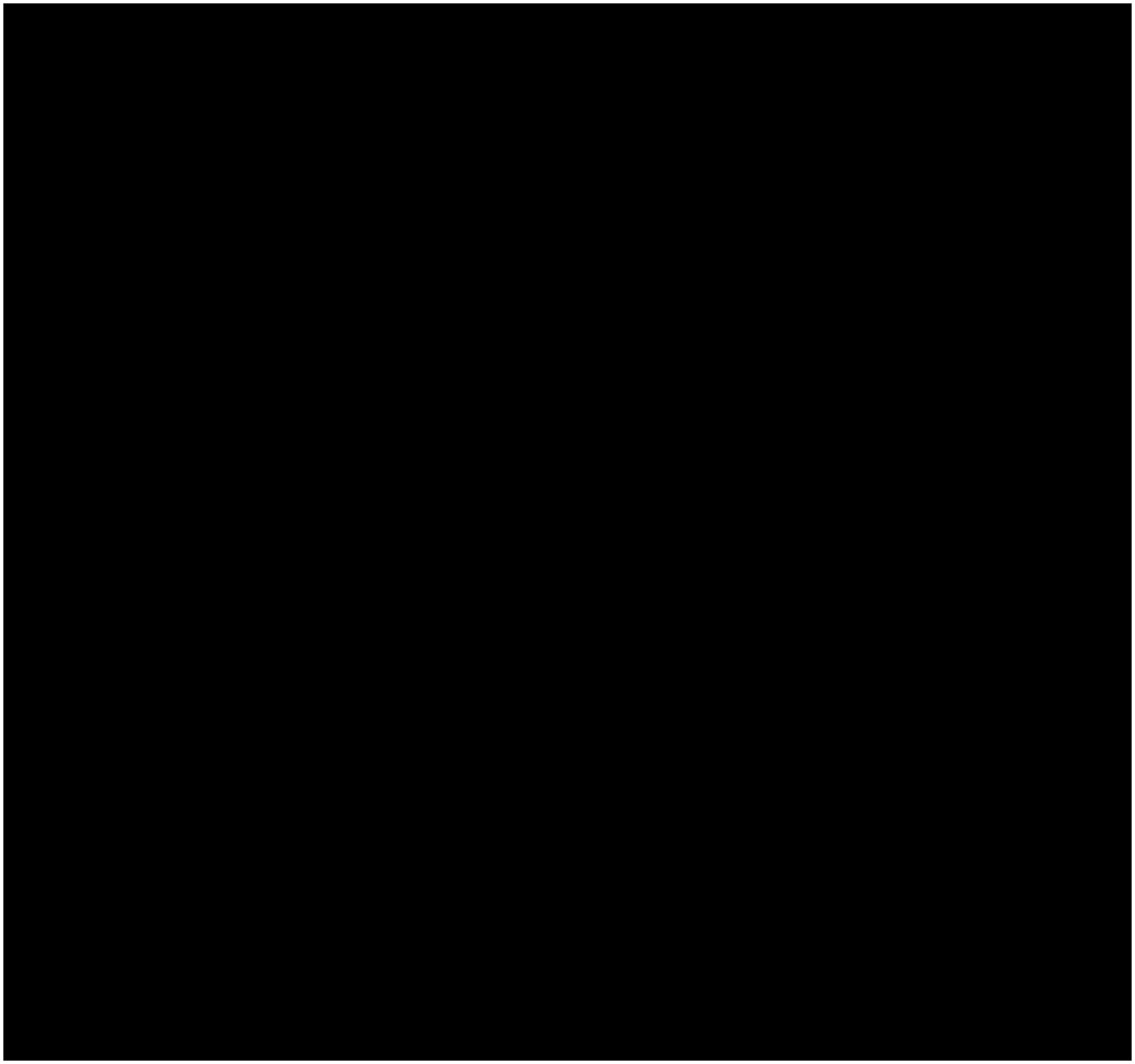
Ramirez A, Lewis G, Eshee L, Sherrod O, Saripkin L, Liu S, et al. Vedolizumab effectiveness associated in children with inflammatory bowel disease (IBD). 2017;65(Supplement 2):S37-8. North American Society for Pediatric Gastroenterology, Hepatology and Nutrition Annual Meeting, Abstract 83.

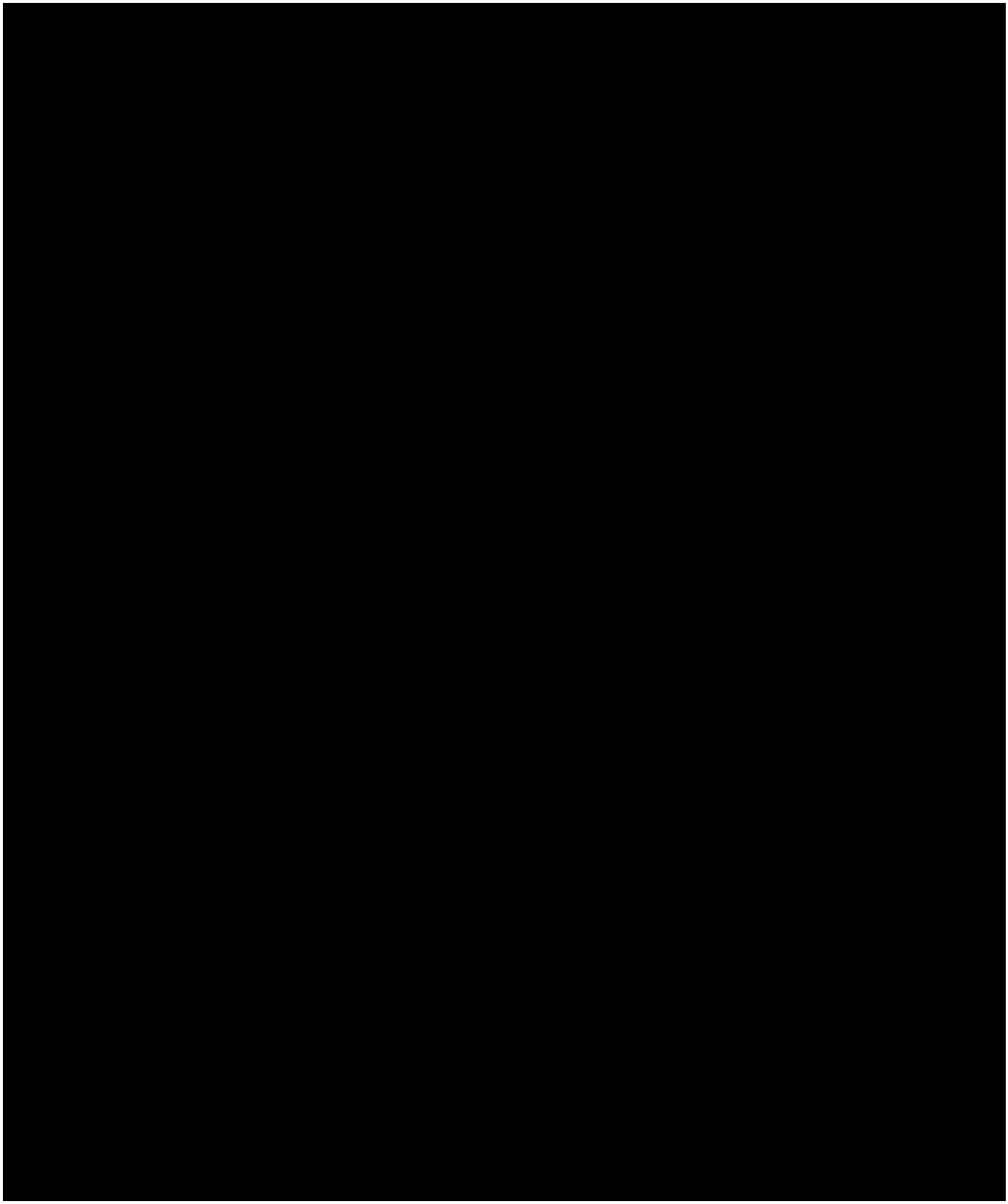
8.0 CHANGES TO PROTOCOL PLANNED ANALYSES

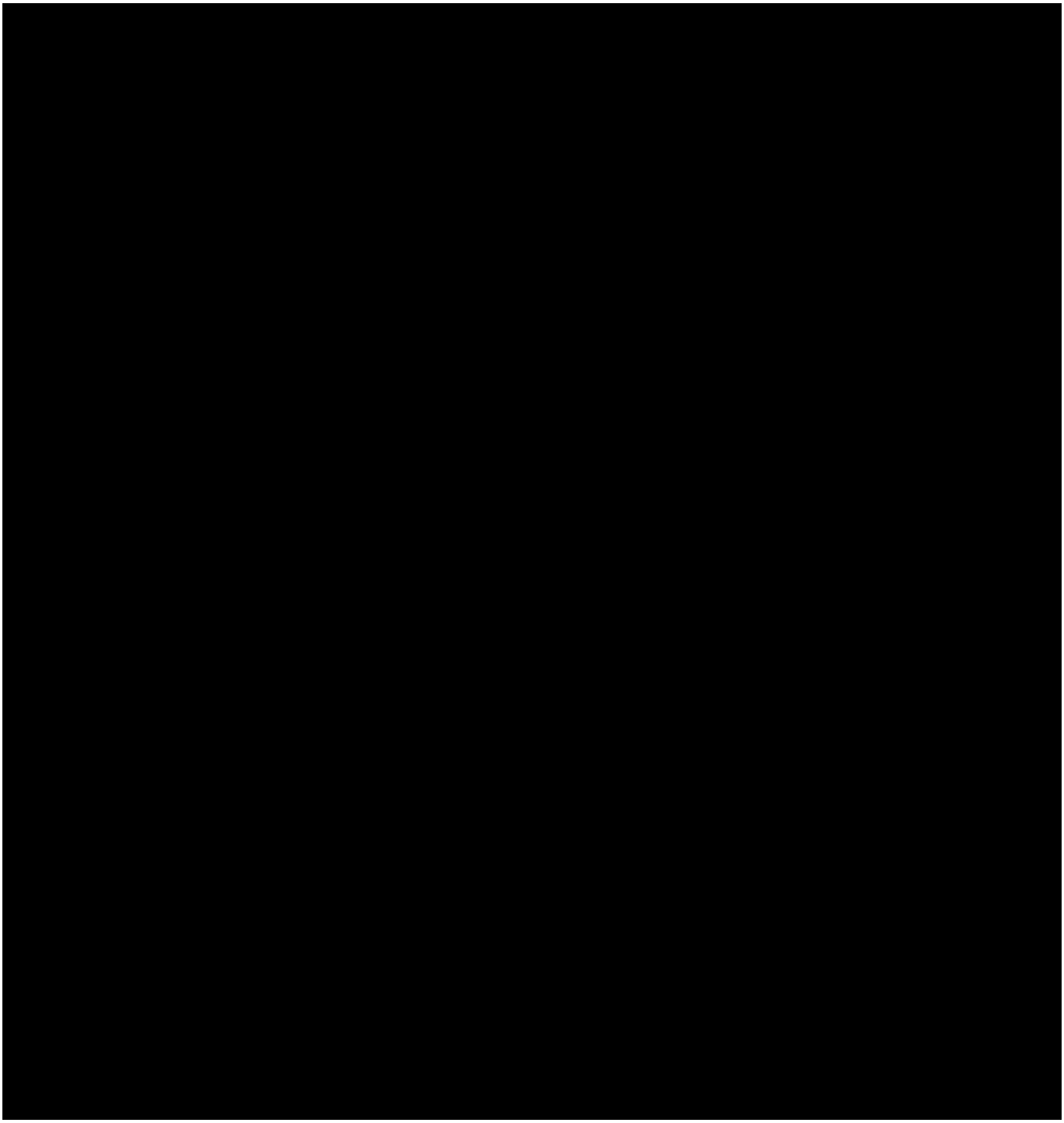
In the Protocol	Modifications Made and Justification
ITT Set definition	Clarified ITT Set definition to account for specific variables
Per-Protocol Set	Renamed as Per-protocol Set – Maintenance (PPS-M) to account for the Week 14 and Week 54 algorithm

9.0 APPENDIX

9.1 Changes From the Previous Version of the SAP







9.2 Data Handling Conventions

9.2.1 General Data Reporting Conventions

For the efficacy analysis, the data across the 3 weight categories (≥30 kg; >15 to <30 kg; 10 to 15 kg) within each dose category (low dose, high dose) will be combined and analyzed by dose (low dose, high dose), unless otherwise specified. This is because PK exposure and response from each planned dose (i.e., low dose, high dose) are expected to be similar across the 3 weight categories (≥30 kg; >15 to <30 kg; 10 to 15 kg). In other words, the high dose in the ≥30 kg weight group is expected to have similar effect as the high dose in the >15 to <30 kg and 10 to 15 kg weight groups. Similarly, the low dose within each weight group (≥30 kg; >15 to <30 kg; 10 to 15 kg) are expected to have similar effect.

9.2.2 Definition of Baseline

Unless otherwise specified, baseline is defined as the last observed non-missing value before the first dose of study drug (utilizing the date and time information). Values measured on Day 1 must be prior to administration of study drug to be classified as baseline.

9.2.3 Definition of Week 14

Week 14 is the randomization day in Maintenance phase.

9.2.4 Definition of Visit Windows

Subjects do not always adhere strictly to the visit timing stated in the protocol. Therefore, the designation of visits will be based on the day of evaluation relative to the start of study drug rather than the nominal visit recorded in the data. Accordingly, the study is divided into continuous, mutually exclusive analysis windows.

The rules provided in Table 9.a Analysis Visit Windows for Safety and Efficacy Data, below, will be used for safety and efficacy data. The lower and upper bounds of each window are the approximate midpoints between the scheduled days for the current visit and its adjacent scheduled visits. The value used in analysis for by-visit summaries is the value within the specified window.

If a patient has more than one measurement within an analysis window, the assessment closest to the target day will be used. In case of ties between observations located on different sides of the target day, the later assessment will be used in analyses. If multiple assessments remain, the scheduled visit measurement will be prioritized for analysis.

Table 9.a Analysis Visit Windows for Safety and Efficacy Data

Analysis Visit	Target Day	Laboratory, Vital Sign, Scores
Baseline	≤ 1 [a]	≤ 1 [a]
Week 2	15	2 – 28
Week 6	43	29 – 56

Table 9.a Analysis Visit Windows for Safety and Efficacy Data

Analysis Visit	Target Day	Laboratory, Vital Sign, Scores
Week 10	71	57 – 84
Week 14	99	85 – 126
Week 22	155	127 – 182
Week 30	211	183 – 238
Week 38	267	239 – 294
Week 46	323	295 – 350
Week 54	379	351- 413
Safety Follow-up Visit		

[a] Baseline will be last observation prior to study drug using date and time.

[b] Safety Follow-up Visit will be based on nominal visit of Safety follow up visit. The data collected during the safety follow up visit should not be allocated to the regular analysis visits up to Week 54.

For data collected during Early Termination visits will be allocated to the regular analysis visits based on the study day.

For PK and AVA data, nominal visits that are within the analysis visit windows will be flagged for analysis. Once a nominal visit falls outside the analysis visit window, the data past that timepoint will not be used for analysis. All PK data with appropriate flags will be in the listings only.

For Mayo scores derivations, the anchor that was found for the primary analysis will be used for sensitivity analysis based on GEMINI approach and sensitivity analysis based on FDA Guidance (2022) approach.

9.3 Analysis Software

SAS Version 9.4 or greater.

9.4 Mayo Score Calculation Worksheet

Complete and Partial Mayo Scoring "Points to Remember"
The Mayo Score is widely used in clinical trials to assess Ulcerative Colitis disease activity. It is a combination of two patient-reported and two physician-determined components. The Partial Mayo Score includes only the Stool Frequency, Rectal Bleeding, and PGA subscores. (Does not include endoscopy)

Sub Scores

Stool Frequency **(Patient)** **0 = Normal number of stools for this patient** **1 = 1 to 2 stools more than normal** **2 = 3 to 4 stools more than normal** **3 = 5 or more stools more than normal**	Stool frequency **WILL:** ➢ Be derived from patient reported diary data in IVRS and will be the average of 3 days prior to visit ➢ Be variable from patient to patient. Instruct patients to set the baseline of "normal" to whatever is "normal" for them. (e.g., A patient normally has 1 stool per day and today has had 4 stools. Therefore the patient has had 3 more than "normal", which yields a value of 2 for that day) ➢ Be defined as the passage of solid or liquid fecal material. Episodes of incontinence proportion. A non-productive trip to the bathroom or the simple passage of gas **DO NOT PROPORTION** as a stool.
Rectal Bleeding **(Patient)** **0 = No blood seen** **1 = Streaks of blood with stool less than half the time** **2 = Obvious blood with stool most of the time** **3 = Blood alone passes**	Rectal bleeding **WILL:** ➢ Be derived from patient reported diary data in IVRS and will be the average of 3 days prior to visit ➢ Represent the most severe bleeding of the day. Hemorrhoidal bleeding **DOES NOT PROPORTION.**
Findings on Endoscopy **(Physician)** **0 = Normal or inactive disease** **1 = Mild disease (erythema, decreased vascular pattern)** **2 = Moderate disease (marked erythema, lack of vascular pattern, friability, erosions)** **3 = Severe disease (spontaneous bleeding, ulceration)**	Findings on Endoscopy **WILL:** ➢ Be documented by photographic evidence ➢ Be classified by the worst affected segment if mucosal appearance varies ➢ Be characterized as follows • Moderate: Bleeds to touch (forceps applied to colonic mucosa for 1 second) • Severe: Bleeds spontaneously ➢ Endoscopy should be performed by the same endoscopist for any given patient

Complete and Partial Mayo Scoring "Points to Remember"

The Mayo Score is widely used in clinical trials to assess Ulcerative Colitis disease activity. It is a combination of two patient-reported and two physician-determined components. The Partial Mayo Score includes only the Stool Frequency, Rectal Bleeding, and PGA subscores. (Does not include endoscopy)

Sub Scores

Physician's Global Assessment **(Physician)** **0 = Normal** **1 = Mild disease** **2 = Moderate disease** **3 = Severe disease**	Physician's Global Assessment **WILL:** ➢ Be based on the patient's overall status on the day of visit ➢ Reflect how the patient is doing at present. Assessment **SHOULD NOT** reflect past disease severity or complexity or the number/kinds of medications the patient is receiving. ➢ Be based on the • Other 3 components of the Mayo score • Patient's recollection of abdominal discomfort and general sense of well-being • Patient's performance status, fecal incontinence, and mood • Physician's observations and physical exam findings ➢ Reflect disease activity, **NOT** disease severity (e.g. **Do not** automatically give a high PGA to patients with pancolitis or severe/complicated disease, or patients requiring multiple medications.)

- Subscores representing the average of 3 days of patient diary data can be obtained from the IVRS subscore report. If calculated manually, **subscores should be rounded to the nearest integer.**
- The Mayo score is equal to the sum of the subscores.

9.5 PUCAI

Time period for evaluation:

- Answers should reflect a daily average of the last 2 days.
- however, if clinical conditions are changing rapidly (e.g., during intense intravenous therapy), the most recent 24 hours should be considered; and
- for patients undergoing colonoscopy, answers should reflect the 2 days before bowel clean out was started.

	Item	Points
1	**Abdominal pain**	
	no pain	0
	pain can be ignored	5
	pain cannot be ignored	10
2	**Rectal bleeding**	
	none	0
	small amount only, in <50% of stools	10
	small amount with most stools	20

	Item	Points
	Large amount	30
3	**Stool consistency of most stools**	
	Formed	0
	Partially formed	5
	Completely unformed	10
4	**Number of stools per 24 hours**	
	0-2	0
	3-5	5
	6-8	10
	>8	15
5	**Nocturnal stools (any episode causing wakening)**	
	No	0
	Yes	10
6	**Activity level**	
	No limitation of activity	0
	Occasional limitation of activity	5
	Severe restricted activity	10
	SCORE	**Total Max 85**

Source: Turner D, Otley AR, Mack D, et al. Development and evaluation of a Pediatric Ulcerative Colitis Activity Index (PUCAI): a prospective multicenter study. Gastroenterology. 2007; 133: 423

9.6 [redacted]

[redacted]

[redacted]

9.7 Adverse Events of Special Interest

Based on the mechanism of action of vedolizumab, certain AESIs have been predefined. The categories of adverse events of special interest, and other planned analyses, are described below.

Events	MedDRA Terms or definitions
Malignancies	SOC: NEOPLASMS BENIGN, MALIGNANT AND UNSPECIFIED (INCL CYSTS AND POLYPS)
Serious Infections [1]	SOC: INFECTIONS AND INFESTATIONS
Infusion Related Reactions	Investigator defined Infusion Related Reactions (as indicated on the AE CRF).
Hypersensitivity Reactions	Anaphylactic/anaphylactoid shock conditions SMQ (broad). Angioedema SMQ (broad). Hypersensitivity SMQ (broad).
PML	Human polyomavirus infection PT JC virus infection PT Leukoencephalopathy PT Progressive multifocal leukoencephalopathy PT JC virus CSF test positive PT Polyomavirus test positive PT JC polyomavirus test positive PT
Liver injury	Cholestasis and jaundice of hepatic origin SMQ (Broad) Hepatic failure, fibrosis and cirrhosis and other liver damage-related conditions SMQ (Broad) Hepatitis, non-infectious SMQ (Broad) Liver related investigations, signs and symptoms SMQ (Narrow) Liver infections SMQ (Broad)

[1] Only Serious events within this SOC are considered AESIs.

9.8 Criteria for Identification of Markedly Abnormal Laboratory Values

Hematology—Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
Hemoglobin	<0.8 × LLN,	>1.2 × ULN
Hematocrit	<0.8 × LLN,	>1.2 × ULN
RBC proportion	<0.8 × LLN,	>1.2 × ULN
WBC proportion	$<2.0 \times 10^3/\mu L$	>1.5 × ULN
Platelet proportion	$<70 \times 10^3/\mu L$	$>600 \times 10^3/\mu L$

RBC=red blood cell, WBC=white blood cell. LLN=lower limit of normal, ULN=upper limit of normal.

Chemistry—Criteria for Markedly Abnormal Values

Parameter	Low Abnormal	High Abnormal
ALT	--	>3x ULN
AST	--	>3x ULN
GGT	--	>3x ULN
Alkaline phosphatase	--	>3x ULN
Total bilirubin	--	>34.2 umol/L
Albumin	<25 g/L	--
Total protein	<0.8x LLN	>1.2x ULN
Creatinine	--	>176.8 umol/L
Sodium	<130 mmol/L	>150 mmol/L
Potassium	<3.0 mmol/L	>6.0 mmol/L
Bicarbonate	<8.0 mmol/L	--
Chloride	<75 mmol/L	>126 mmol/L
Calcium	<1.50 mmol/L	>3.25 mmol/L
Glucose	≤2.8 mmol/L	≥20 mmol/L
Phosphorous	<0.52 mmol/L	>2.10 mmol/L
CPK	--	>5x ULN

ALT=alanine aminotransferase, AST=aspartate aminotransferase, GGT=□-glutamyl transferase, CPK=creatine phosphokinase, LLN=lower limit of normal, ULN=upper limit of normal.

9.9 Criteria for Markedly Abnormal Values for Vital Signs

Parameter	Unit	Lower Criteria	Upper Criteria
Pulse	bpm	<50	>120
Systolic blood pressure	mm Hg	<85	>180
Diastolic blood pressure	mm Hg	<50	>110
Body temperature	ºC	< 35.6	>37.7
	ºF	<96.1	>99.9

9.10 Z-Score Derivations

Official and validated SAS programs created by Centers for Disease Control and Prevention (CDC) will be used to calculate the Z-scores (standard deviations) for a child's sex and age (up to 20 years of age) for BMI, weight, and height based on the CDC growth charts for children age 2 years and older and the WHO growth charts for infants and children < 2 years of age. For more information on the CDC SAS programs, see

http://www.cdc.gov/growthcharts/computer_programs.htm.

Z-scores are calculated as the formula below:

Z-score= [((observed value / M) ^ L) – 1] / (S * L)

In which 'observed value' is the child's height, weight or derived BMI. The L, M, and S values vary according to the child's sex and age. For more information on the LMS method, see http://www.ncbi.nlm.nih.gov/pmc/articles/PMC27365/.

Signature Page for Vedolizumab-3024 6-1-9-6 3024 Statistical Analysis Plan Amend Title:

Approval Task	[redacted] Statistics 10-Jun-2025 18:59:05 GMT+0000

Document Number: TDN-000557004 v1.0