# STATISTICAL ANALYSIS PLAN
# FOR
# PLACEBO-CONTROLLED PERIOD WEEK 12 ANALYSIS
# AND
# PLACEBO-CONTROLLED PERIOD FINAL ANALYSIS

**FINAL**
**(Version 1.0)**
**11 April 2022**

Protocol Number: GB004-2101
(Version 3.0, 15 December 2021)

## A Phase 2, randomized, double-blind, placebo-controlled, multi-center study to evaluate GB004 in adult subjects with mild-to-moderate active ulcerative colitis

SPONSORED BY

GB004, Inc., a wholly owned subsidiary of Gossamer Bio, Inc.

PREPARED BY

Gossamer Bio Services, Inc.

# VERSION CONTROL

| Version Number | Date | Comments/Changes |
|---|---|---|
| 1.0 | 11 April 2022 | Final |

[redacted]

# TABLE OF CONTENTS

# LIST OF TABLES

## LIST OF ABBREVIATIONS

| Abbreviation | Definition |
|---|---|
| 5-ASA | 5-aminosalicylate |
| AE | adverse event |
| ALT | alanine aminotransferase |
| ANCOVA | analysis of covariance |
| AST | aspartate aminotransferase |
| ATC | anatomical therapeutic chemical |
| AUC | area under the drug concentration versus time curve |
| BID | twice per day |
| BMI | body mass index |
| bpm | beats per minute |
| CI | confidence interval |
| CKD-EPI | Chronic Kidney Disease Epidemiology Collaboration |
| $C_{max}$ | maximum observed concentration |
| CMH | Cochran-Mantel-Haenszel |
| COVID-19 | Coronavirus Disease 2019 |
| CRP | C-reactive protein |
| CSR | clinical study report |
| DRC | data review committee |
| ECG | electrocardiogram |
| eCRF | electronic case report form |
| eDiary | electronic diary |
| eDISH | evaluation of drug-induced serious hepatotoxicity |
| eGFR | estimated glomerular filtration rate |
| EMA | European Medicines Agency |
| EPO | erythropoietin |
| FDA | Food and Drug Administration |
| GGT | gamma-glutamyl transferase |
| Hg | mercury |
| HIF-1α | Hypoxia-inducible factor 1 alpha |
| ICE | intercurrent event |
| ICH | International Conference on Harmonisation |

| Abbreviation | Definition |
|---|---|
| IP | investigational product |
| IRT | interactive response technology |
| ITT | intent-to-treat |
| LS | least squares |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MAR | missing at random |
| MI | multiple imputation |
| MNAR | missing not at random |
| MPO | myeloperoxidase |
| NRI | non-responder imputation |
| OLE | Open-Label Extension |
| OR | odds ratio |
| PCP | Placebo-Controlled Period |
| PGA | Physician's Global Assessment |
| PK | pharmacokinetic(s) |
| PT | preferred term |
| QD | once daily |
| QTcF | QT interval corrected for heart rate using Fridericia's formula |
| RBS | rectal bleeding subscore |
| RHI | Robarts Histopathology Index |
| SAE | serious adverse event |
| SAP | statistical analysis plan |
| SCr | serum creatinine |
| SD | standard deviation |
| SE | standard error |
| SFS | stool frequency subscore |
| SoA | schedule of activities |
| SOC | system organ class |
| TEAE | treatment-emergent adverse event |
| $T_{max}$ | time of occurrence of maximum observed concentration |
| UC | ulcerative colitis |
| UC-100 | Ulcerative Colitis 100 index |

| Abbreviation | Definition |
|---|---|
| ULN | upper limit of normal |
| VEGF | vascular endothelial growth factor |
| WHO | World Health Organization |

# 1. INTRODUCTION

This statistical analysis plan (SAP) provides a comprehensive and detailed description of the statistical methodology to be used for analysis and reporting of study results for the Placebo-Controlled Period (PCP) Week 12 Analysis and PCP Final Analysis of Study GB004-2101 and is based on protocol version 3.0 dated 15 December 2021.

This SAP has been finalized and approved prior to unblinding of the Sponsor for the PCP Week 12 Analysis. A separate SAP describes the statistical methodology to be used for analysis and reporting of study results for the Open-Label Extension (OLE) Final Analysis. The 3 separate formal statistical analyses to be performed for the study (PCP Week 12 Analysis, PCP Final Analysis, and OLE Final Analysis) are described in Protocol Section 9.4.

Results to be reported will include summaries of subject disposition, demographic and baseline characteristics, major protocol deviations, prior and concomitant medications, medical history, duration of study treatment exposure and compliance, and primary, secondary, safety, and select exploratory endpoints.

If additional analyses are required to supplement the planned analyses described in this SAP, they will be completed and identified in the clinical study report (CSR) as post hoc.

This SAP was written with due consideration of the recommendations outlined in the most recent International Conference on Harmonisation (ICH) E9 Guideline (ICH E9, 1998) entitled Guidance for Industry E9 Statistical Principles for Clinical Trials, the most recent ICH E9(R1) Guideline (ICH E9(R1), 2019) entitled Addendum on Estimands and Sensitivity Analyses in Clinical Trials to the Guideline on Statistical Principles for Clinical Trials, and the most recent ICH E3 Guideline (ICH E3, 1995) entitled Guidance for Industry Structure and Content of Clinical Study Reports.

In this document, investigational product (IP) and study treatment have the same meaning and refer to placebo, GB004 480 mg once daily (QD), and GB004 480 mg twice per day (BID).

# 2. PROTOCOL SUMMARY

This is a Phase 2, randomized, double-blind, placebo-controlled, multi-center study to evaluate GB004 in adult subjects with mild-to-moderate active ulcerative colitis (UC). The study consists of two parts, a 36-week placebo-controlled period and a 24-week open-label extension period, and will assess the efficacy and safety of GB004 when added to background UC therapy of 5-aminosalicylate (5-ASA) with or without systemic corticosteroids.

## 2.1. Placebo-Controlled Period Objectives and Endpoints

Table 1 below, from Protocol Section 3.1, lists the PCP objectives and endpoints.

**Table 1: Placebo-Controlled Period Objectives and Endpoints**

| Objectives | Endpoints |
|---|---|
| **Primary** | |
| • To evaluate the effect of GB004 compared to placebo on clinical remission at PCP Week 12 | • Proportion of subjects with clinical remission at PCP Week 12, defined as a Modified Mayo score ≤ 2, with a Rectal bleeding subscore of 0, Stool frequency subscore of 0 or 1 (with a ≥ 1 point decrease from baseline), and Endoscopic subscore of 0 or 1 |
| **Secondary** | |
| • To evaluate the effect of GB004 on clinical response, histologic remission, endoscopic improvement and mucosal healing at PCP Week 12 | • Proportion of subjects with clinical response at PCP Week 12, defined as reduction in the Modified Mayo score of ≥ 2 points and ≥ 35 percent reduction from baseline, including a decrease in Rectal bleeding subscore of ≥1 or absolute Rectal bleeding subscore of ≤ 1<br>• Proportion of subjects with histologic remission at PCP Week 12, defined as Robarts Histopathology Index (RHI) ≤ 3 with lamina propria neutrophils subscore = 0 and neutrophils in epithelium subscore = 0<br>• Proportion of subjects with endoscopic improvement at PCP Week 12, defined as endoscopic subscore of 0 or 1<br>• Proportion of subjects with mucosal healing at PCP Week 12, defined as endoscopic improvement and histologic remission |
| • To evaluate the effect of GB004 on clinical remission, clinical response, histologic remission, endoscopic improvement, and mucosal healing at PCP Week 36 | • Proportion of subjects with clinical remission at PCP Week 36<br>• Proportion of subjects with clinical response at PCP Week 36<br>• Proportion of subjects with histologic remission at PCP Week 36<br>• Proportion of subjects with endoscopic improvement at PCP Week 36<br>• Proportion of subjects with mucosal healing at PCP Week 36 |
| **Safety** | |
| • To evaluate the safety and tolerability of GB004 | • Incidence of treatment emergent adverse events (TEAEs) |

| Objectives | Endpoints |
|---|---|
| **Exploratory** | |
| • To evaluate other safety outcomes | • Change from Baseline in laboratory, vital signs, and electrocardiogram (ECG) parameters |
| • To explore other measures of the effect of GB004 on disease activity over time | • Proportion of subjects with resolution of Rectal bleeding<br>• Proportion of subjects with histologic response<br>• Proportion of subjects with Geboes score ≤ 2<br>• Proportion of subjects with steroid-free clinical remission at PCP Week 36<br>• Proportion of subjects with clinical remission by Mayo score and resolution of rectal bleeding<br>• Proportion of subjects with clinical remission by Mayo score<br>• Proportion of subjects with clinical response by Mayo score<br>• Proportion of subjects with Modified Mayo score of ≤ 2, with no individual subscore > 1<br>• Proportion of subjects with disease clearance<br>• Proportion of subjects with symptomatic remission<br>• Proportion of subjects with modified symptomatic remission<br>• Proportion of subjects with Partial Mayo remission<br>• Proportion of subjects with reduction of ≥ 2 points from baseline in partial Mayo score<br>• Proportion of subjects with a 6-point Mayo score clinical remission<br>• Change from Baseline in Modified Mayo score<br>• Change from Baseline in Mayo score<br>• Change from Baseline in Partial Mayo score over time<br>• Change from Baseline in Mayo subscores over time<br>• Change from Baseline in RHI score<br>• Change from Baseline in the UC-100 Index |
| • To explore the pharmacokinetics (PK) of GB004 | • Plasma and colon tissue concentrations of GB004 |

| Objectives | Endpoints |
|---|---|
| • To explore target engagement and pharmacodynamic response | • Change from baseline in tissue and blood markers of target engagement, eg, HIF-1α, EGLN1-2-3, etc.<br>• Change from baseline in markers of pharmacodynamics, eg, fecal calprotectin, fecal lactoferrin, and neutrophil activity (myeloperoxidase)<br>• Change from baseline in tissue, blood, and stool in downstream genes and proteins, eg, CAIX, MPO, c-reactive protein (CRP), TJP, CLDN1, etc. as a function of exposure ($C_{max}$, $T_{max}$, AUC), efficacy endpoints (eg, clinical remission, clinical response), and/or safety endpoints |
| • To explore baseline markers with response to treatment with GB004 | • Relationship between tissue, blood, and stool baseline markers (eg, HIF-1α, EGLN1-2-3, calprotectin, etc.) and response to treatment, as measured by efficacy endpoints (eg, clinical remission, clinical response), and/or safety endpoints |
| • To explore pharmacogenetics | • Effect of naturally occurring genetic variation on the efficacy (eg, clinical remission, clinical response), safety, and/or PK profile ($C_{max}$, $T_{max}$, AUC) and tissue concentrations associated with treatment with GB004 |

The last three Exploratory objectives listed in Table 1 above and their corresponding endpoints will not be addressed in this SAP, with the exception of the endpoints for change from baseline in fecal calprotectin, fecal lactoferrin, and CRP.

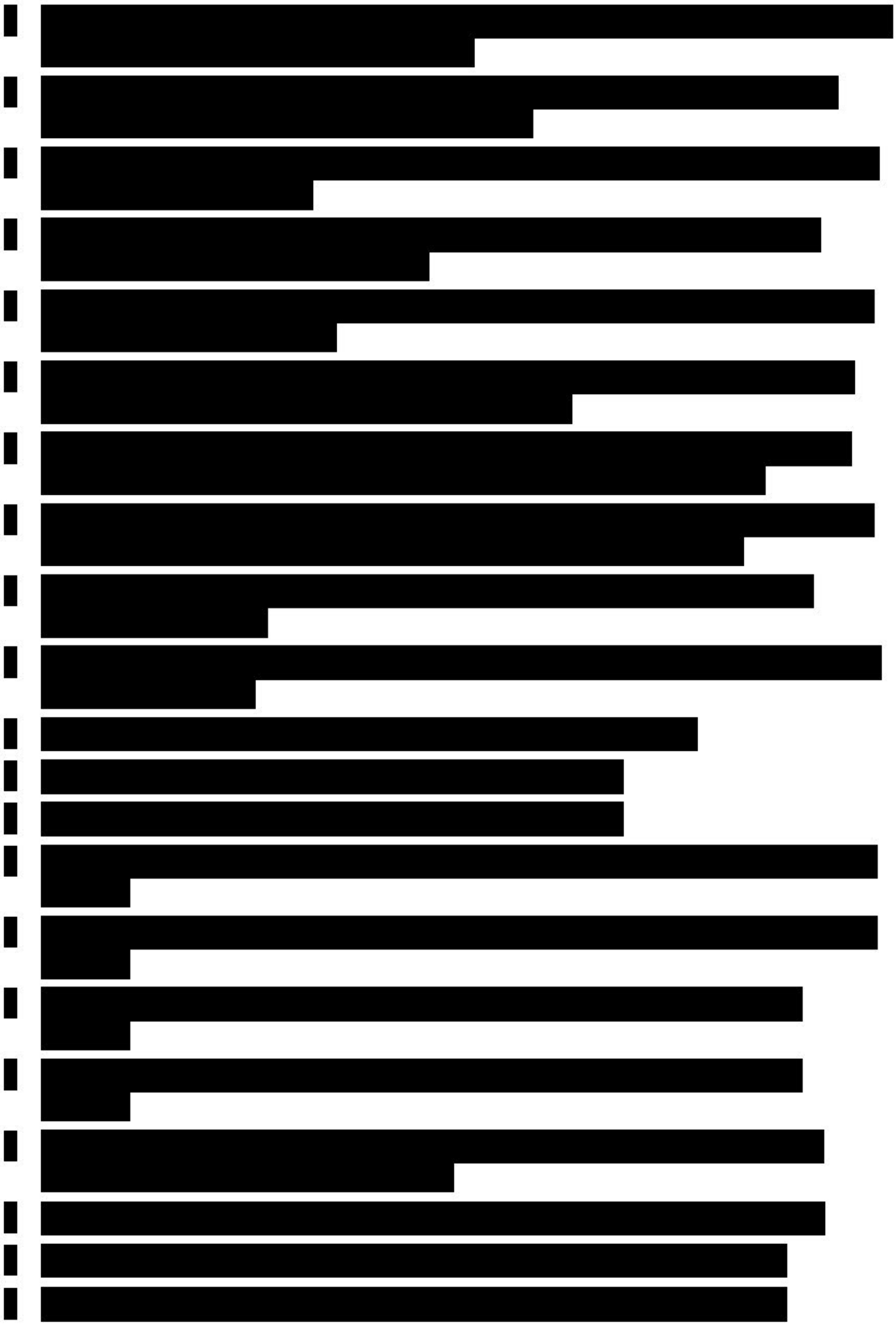
## 2.2. Efficacy Endpoint Definitions

Table 2 below, from Protocol Section 3.3, defines the PCP efficacy endpoints.

**Table 2: Placebo-Controlled Period Efficacy Endpoint Definitions**

| Endpoint | Definitions |
|---|---|
| Clinical remission | Modified Mayo score ≤ 2, with a Rectal bleeding subscore of 0, Stool frequency subscore of 0 or 1 (with ≥1 point decrease from baseline), and Endoscopic subscore of 0 or 1 |
| Clinical response | Reduction in the Modified Mayo score of ≥ 2 points and ≥ 35% from baseline, including a decrease in Rectal bleeding subscore of ≥1 or absolute Rectal bleeding subscore of ≤ 1 |
| Histologic remission | RHI ≤ 3 with lamina propria neutrophils subscore = 0 and neutrophils in epithelium subscore = 0 |
| Endoscopic improvement | Endoscopic subscore 0 or 1 |
| Mucosal healing | Endoscopic improvement and histologic remission |
| Resolution of Rectal bleeding | Rectal bleeding subscore = 0 |
| Histologic response | Decrease from baseline in RHI of ≥ 7 |
| Steroid-free remission | Clinical remission with no systemic corticosteroid use among subjects who were using systemic corticosteroids at baseline |
| Clinical remission by Mayo score | Mayo score ≤ 2, with no Mayo individual subscores >1 |
| Clinical response by Mayo score | Reduction in the Mayo score of ≥ 3 points and ≥ 30% from baseline, including a decrease in Rectal bleeding subscore of ≥ 1 or absolute Rectal bleeding subscore of ≤1 |
| Disease clearance | Clinical remission and histological remission |
| Symptomatic remission | Mayo score of ≤ 2, with no individual subscore > 1 and both Rectal bleeding and Stool frequency subscores of 0 |
| Modified symptomatic remission | Endoscopic improvement, resolution of Rectal bleeding and Stool frequency subscore of 0 |
| Partial Mayo remission | Partial Mayo score of ≤ 2 with no individual subscore > 1 |
| 6-point Mayo score clinical remission | 6- point Mayo score < 2 with a Rectal bleeding subscore of 0 and Stool frequency subscore of 0 or 1 (with ≥ 1 point decrease from baseline) |

[redacted]



## 2.3. Overall Study Design and Plan

This is a 2-part study in adult subjects with mild-to-moderate active UC who have disease activity despite treatment with 5-ASA with or without systemic corticosteroids:

- The PCP is a randomized, placebo-controlled, multi-center, 36-week study evaluating the efficacy, safety, tolerability and PK of 2 dose regimens of GB004; the initial 12 weeks of treatment will be double-blind.
- The OLE is an open-label, multi-center, 24-week study evaluating the safety and tolerability of GB004 twice per day (BID) dose regimen.

Subjects who complete the Week 36 visit on IP and subjects who meet the predefined UC Disease Activity criteria (defined in Protocol Section 5.3 and Protocol Appendix 8) at or after PCP Week 12 visit and prior to PCP Week 36 visit on IP can enter the OLE.

All subjects participating in the study will be required to maintain a stable dose of 5-ASA throughout the study. Subjects on systemic corticosteroids must remain on a stable dose during the initial 12 weeks of the PCP. After the PCP Week 12 visit, a standardized taper from systemic corticosteroids may be attempted at any time.

All subjects will attend a Follow-up visit at the clinic 4 weeks after last dose of IP.

A data review committee (DRC) will periodically convene to review unblinded overall safety and emerging efficacy results.

A schematic of the study design and the schedule of activities (SoA) are provided in Protocol Sections 1.2 and 1.3, respectively.

### 2.3.1. Placebo-Controlled Period Study Design and Plan

After signing an informed consent form, subjects will be screened for study eligibility over a Screening period of up to 5 weeks. During the Screening period, subjects will capture stool frequency and rectal bleeding symptoms in a provided electronic diary (eDiary) on a daily basis. Flexible sigmoidoscopy or colonoscopy will be performed at screening. Subjects not meeting the eligibility criteria will be deemed screen failures and will not continue participation in the study.

On Day 1/Week 0, eligible subjects will be randomized 1:1:1 to receive the following treatments for 36 weeks:

- GB004 480 mg BID
- GB004 480 mg once daily (QD)

- Placebo

Randomization will be stratified by systemic corticosteroid use at baseline (yes/no).

All subjects will dose IP BID to maintain the blind.

The first dose of IP will be administered in the clinic on Day 1/Week 0 with food. All subsequent doses will be taken at home with food, with the exception of the morning doses at the Week 2, 4, 8, 12, and 36 visits where IP will be administered in the clinic, with food, after predose blood collection.

After initiation of IP on Day 1/Week 0, subjects will return to the clinic and will be evaluated as specified in the Schedule of Activities (SoA) of the protocol. A flexible sigmoidoscopy with biopsies will be performed at the PCP Week 12 and Week 36 visits (and at the Unscheduled UC Disease Activity Criteria Assessment visit(s) and at the Early Withdrawal from Study visit, if applicable).

Subjects who permanently discontinue PCP IP without enrolling in the OLE will be encouraged to continue in the PCP and will complete the Early Discontinuation of IP visit at the time of IP discontinuation and then complete the remaining PCP study visits per the SoA. At subsequent visits, all study procedures will be completed per the SoA excluding dispensation/return and accountability of IP.

Subjects who withdraw from the PCP without enrolling in the OLE, regardless of the reason, will be requested to return to the clinic to complete the Early Withdrawal from Study visit and will be asked to return for a Follow-up visit approximately 4 weeks after their last dose of IP to assess safety.

Total duration for PCP participation per subject is up to 45 weeks, including a Screening period of up to 5 weeks, a treatment period of up to 36 weeks, and a 4-week safety Follow-up period.

Subjects may participate in the 24-week OLE if in the opinion of the Investigator they have been compliant with study procedures and:

- Completed the Week 36 visit on IP; or
- Met the predefined UC Disease Activity criteria any time after PCP Week 12 visit and prior to PCP Week 36 visit on IP.

## 2.4. Study Population

The study population consists of adult subjects with mild-to-moderate active UC who have disease activity despite treatment with 5-ASA with or without prednisone (or equivalent), beclomethasone, budesonide, or budesonide multi-matrix. Mild-to-moderate UC is defined as a Mayo score of 5-10, inclusive, with an Endoscopic subscore of $\geq 2$, a stool frequency subscore $\geq 1$, and a Rectal bleeding subscore $\geq 1$.

## 2.5. Randomization, Stratification, and Blinding

### 2.5.1. Randomization and Stratification

The study will randomize approximately 195 subjects, with approximately 65 subjects per treatment group, in a 1:1:1 ratio to GB004 480 mg BID, GB004 480 mg QD, and Placebo.

Randomization will be performed utilizing permuted block randomization via an interactive response technology (IRT) system. Randomization will be stratified by systemic corticosteroid use at baseline (yes/no).

### 2.5.2. Blinding

The PCP is considered double-blind through PCP Week 12, as the study will be unblinded to the Sponsor at the time of the PCP Week 12 Analysis. The Sponsor (with exceptions described below) will be blinded to individual subject treatment assignments until the PCP Week 12 Analysis. Subjects, investigators and site personnel, personnel performing central endoscopic and histologic evaluations, and personnel performing central laboratory and ECG evaluations will remain blinded to individual subject treatment assignments throughout the PCP.

Sponsor (or designee) personnel will have access to unblinded individual subject treatment assignments prior to unblinding of the Sponsor at the time of the PCP Week 12 Analysis for the purposes of study-required activities, including management of IP inventory, production of summaries of data for DRC review, performance of bioanalytical analysis of PK and gene/protein concentrations, and performance of population pharmacokinetic modeling, for which subject level treatment assignment data will be provided to clinical pharmacology personnel performing the modeling using only pseudonymized (also known as dummy) subject identifiers. These personnel will not be directly involved in the conduct of the study. The DRC, comprised of Sponsor representatives not directly involved in the conduct of the study and external expert physicians, will have access to unblinded study results and data during the course of the study.

## 2.6. Sample Size Determination

A total sample size of approximately 195 subjects (approximately 65 per treatment group, randomized 1:1:1, with stratification by systemic corticosteroid use at baseline [yes/no]) is estimated to provide approximately 80% power to detect a difference of 20.4% between each GB004 treatment group and placebo for the primary endpoint of proportion of subjects with clinical remission at PCP Week 12 based on a chi-squared test at an 0.050 two-sided level of significance. This assumes the proportion of subjects with clinical remission at PCP Week 12 is 10% in the placebo group and 30.4% in each GB004 treatment group and a dropout rate by PCP Week 12 of approximately 8%. Subjects with missing clinical remission status at PCP Week 12 will be considered as having not met the primary endpoint.

Using the same sample size assumptions as above for the primary endpoint of proportion of subjects with clinical remission at PCP Week 12 (ie, a difference of 20.4% between each GB004 treatment group and placebo, with a 10% proportion in the placebo group), there will be approximately 80% power for the secondary endpoint of the proportion of subjects with clinical remission at PCP Week 36 based on a chi-squared test at an 0.050 two-sided level of

significance. Subjects with missing clinical remission status at PCP Week 36 will be considered as having not met the endpoint.

# 3. GENERAL ANALYSIS AND REPORTING CONVENTIONS

Three separate formal statistical analyses will be performed as described in Protocol Section 9.4:

- The PCP Week 12 Analysis will be conducted after all subjects either complete the PCP Week 12 visit or withdraw from the study prior to completion of the PCP Week 12 visit. This analysis will evaluate endpoints at or up through PCP Week 12.
- The PCP Final Analysis will be conducted after all subjects either complete the PCP or withdraw from the study prior to completion of the PCP; This analysis will evaluate all PCP endpoints.
- The OLE Final Analysis will be conducted after all subjects either complete the OLE period or withdraw from the study prior to completion of the OLE period. This analysis will evaluate all OLE endpoints.

This SAP governs the PCP Week 12 Analysis and the PCP Final Analysis.

All analyses will be performed using SAS® System (SAS Institute Inc., Cary, NC) version 9.4 or later.

By-subject listings described in this SAP will be produced for the PCP Final Analysis only.

Categorical variables (ie, those with discrete categories or levels, whether nominal or ordinal) will generally be summarized using descriptive statistics of counts (n) and percentages (%) and will be presented in the format "n (xx.x)". To ensure completeness, summaries for categorical variables may include all categories, even if no subjects had a value in a particular category. Additionally, for missing data, a category of "Missing" will be presented as needed. Percentages will be rounded to one decimal place. If a count is zero, no percentage will be shown, and if a percentage is 100%, 100.0% will be shown.

Continuous variables will generally be summarized using descriptive statistics: number of subjects with non-missing data (n), mean, standard deviation (SD), median, 25th and 75th percentiles, minimum, and maximum. Descriptive statistics on select measures may also include stool standard errors (SEs) and two-sided, 95% confidence intervals (CIs). In general, rounding rules for reporting continuous descriptive summary statistics are as follows:

- If the original values have 0 or 1 decimal places: mean, median, 25th and $75^{th}$ percentiles will be reported to one more decimal place than the original values, and SD and SE will be reported to 2 more decimal places than the original values.
- If the original values have 2 or more decimal places: mean, median, 25th and 75th percentiles, SD, and SE will all be reported to 3 decimal places.

Original values include both reported raw data values and derived values. In the case of derived values, an appropriate determination for precision will be made based on the reported raw data values used for derivation. Minimum and maximum will always be reported to the same number of decimal places as the original values, up to a maximum of 3 decimal places. Percent change

from baseline will be reported to 1 more decimal place than change from baseline, up to a maximum of 3 decimal places.

Efficacy results will be summarized by various measures, including, but not limited to, absolute differences in proportions, odds ratios (ORs), least squares (LS) means, differences in LS means, SEs, p-values, and two-sided, 95% CIs. Absolute differences in proportions and corresponding CIs will be reported to 1 decimal place. ORs and their corresponding CIs will be reported to 3 decimal places. Least squares means and differences in LS means and their corresponding CIs and SEs will be reported per the rules above, with the CIs following the same rounding rules as the corresponding means. P-values will be reported to 4 decimal places, with values less than 0.0001 displayed as < 0.0001 and values greater than 0.9999 displayed as > 0.9999. All statistical hypothesis testing for efficacy endpoints will be at an 0.050 two-sided level of significance.

Summaries of subject disposition, major protocol deviations, demographic and other baseline characteristics, UC treatment history, and concomitant UC background medications at baseline will be presented by randomized treatment group and will also include an overall column combining all treatment groups. Summaries of medical history and prior medications will be presented by actual treatment group and will include an overall column combining all treatment groups. Summaries of concomitant medications, duration of study treatment, compliance, prescribed dose modifications, and time on study will be presented by actual treatment group and will also include a total GB004 column combining both GB004 treatment groups. Efficacy tables will be presented by randomized treatment group only. Safety tables will be presented by actual treatment group and, in general, will also include a total GB004 column combining both GB004 treatment groups.

When appropriate, sorting of non-ordinal categorical variables in tables will be based on decreasing frequency within the overall column. In general, adverse event (AE) tables that are displayed by system organ class (SOC) and preferred term (PT) will be sorted by the internationally agreed order for SOC, and by decreasing incidence and then alphabetically for PT within SOC in the total GB004 column. Otherwise, AE tables, when applicable, will be sorted by decreasing incidence within the total GB004 column.

In general, assessments at the PCP Day 1 visit are to be performed prior to the first dose of study treatment at the PCP Day 1 visit. Therefore, the baseline value is defined as the last non-missing value on or before the first dose of PCP study treatment, unless otherwise specified.

For analyses of change from baseline and percent change from baseline parameters, only subjects with a baseline value and at least one post-baseline value will be included. For maximum post-baseline analyses, only subjects with at least one post-baseline value will be included. Analyses of shift from baseline, outliers, and abnormalities will generally be based on the number of subjects at risk.

Values with “<” or “>” signs will be analyzed without the signs in tables and figures. In by-subject data listings, values will be reported as collected with the sign.

Dates in by-subject data listings will be displayed as yyyy-mm-dd (eg, 2021-01-24). In general, by-subject data listings will be sorted by treatment group, subject number, and study day.

## 3.1. Determination of Data for Inclusion in Placebo-Controlled Period Analyses

In general, statistical analyses described in this SAP will be performed separately for the PCP Week 12 Analysis and for the PCP Final Analysis. The PCP Week 12 Analysis is intended to analyze endpoints at or up through PCP Week 12, and the PCP Final Analysis is intended to evaluate all PCP endpoints. Therefore, a determination of study data to be included in the PCP Week 12 Analysis and the PCP Final Analysis is required.

### 3.1.1. Placebo-Controlled Period Week 12 Analysis

For subjects who consent only to version 1.0 of the protocol, which includes only a 12-week placebo-controlled treatment period, all data reported will be included in the PCP Week 12 Analysis.

For subjects who consent to version 2.0 of the protocol, which includes a 36-week placebo-controlled treatment period and a 24-week OLE treatment period, data to be included in the PCP Week 12 Analysis will generally be dependent upon whether a subject continues in the PCP beyond PCP Week 12 as follows:

- For subjects who withdraw from the PCP prior to PCP Week 12 and subjects who withdraw from the PCP at PCP Week 12 without transitioning into OLE, all data reported, including data from the PCP Follow-up visit, will be included in the PCP Week 12 Analysis.
- For subjects who withdraw from the PCP and transition into the OLE at PCP Week 12, all data up through the date and time (where ascertainable) of first dose of OLE study treatment will be included in the PCP Week 12 Analysis.
- For subjects who continue in the PCP beyond PCP Week 12, all data reported up through and including PCP Week 12 will be included in the PCP Week 12 Analysis.

### 3.1.2. Placebo-Controlled Period Final Analysis

For subjects who consent only to version 1.0 of the protocol, all data reported will be included in the PCP Final Analysis.

For subjects who consent to version 2.0 of the protocol, data to be included in the PCP Final Analysis will generally be dependent on whether a subject is dosed in the OLE as follows:

- For subjects not dosed in the OLE, all data reported will be included in the PCP Final Analysis.
- For subjects dosed in the OLE, all data up through the date and time (where ascertainable) of first dose of OLE study treatment will be included in the PCP Final Analysis.

## 3.2. Analysis Visit Windowing

Subjects do not always strictly adhere to the visit schedule timing in the protocol. Therefore, the designation of analysis visits (or timepoints) will generally be based on the actual day of evaluation relative to the date of first dose of PCP study treatment (Day 1), rather than the

nominal visit, for analyses conducted by visit or analyses conducted at a visit, unless otherwise stated.

Mutually exclusive analysis visit windows containing no gaps will generally be utilized to assign analysis visits corresponding to scheduled post-baseline visits specified in the protocol and will be based on data from both scheduled and unscheduled visits/assessments. Analysis visits will be assigned by using a windowing scheme as described below.

The target day for analysis visits for the PCP Week 2, 4, 8, 12, 18, 24, 30 and 36 visits is the week number multiplied by 7 plus 1 (eg, target day for the PCP Week 2 analysis visit is Day 15). The protocol-specified Follow-up visit in the PCP will have an analysis visit designation of Week 40 (Follow-up), and therefore, the week number for the Follow-up visit will be considered to be Week 40 for analysis visit windowing (ie, target day of 281).

The upper bound of the baseline analysis visit window is Day 1 (up to and including the time of first dose of PCP study treatment for assessments with time collected), and the lower bound of the first post-baseline visit window is Day 1 after the time of first dose of PCP study treatment (for assessments with time collected) or Day 2 (for assessments without time collected). For all other lower and upper bounds of analysis visit windows, with the exceptions noted below, windows will end at the midpoint between scheduled visit timepoints, with the midpoint itself assigned to the latter visit window.

The upper bound of the PCP Week 12 analysis visit window for assessments collected per the protocol SoA at both the PCP Week 12 and PCP Week 18 visits and for the PCP Week 12 flexible sigmoidoscopy with biopsy will be Day 107 (allowing for a 22-day upper analysis visit window after the target day of Day 85). This upper analysis visit window incorporates the protocol-specified 5-day visit window for the PCP Week 12 visit, the additional protocol-specified 14-day allotment to allow for any delay in endoscopy assessment due to a pandemic (eg, Coronavirus Disease 2019 pandemic) or other global health emergency, and the protocol-specified requirement that the PCP Week 12 visit be conducted in-clinic within 3 days post-endoscopy.

Similarly, the upper bound of the PCP Week 36 analysis visit window for assessments collected per the protocol SoA at both the PCP Week 36 and PCP Follow-up visits and for the PCP Week 36 flexible sigmoidoscopy with biopsy will be Day 275 (allowing for a 22-day upper analysis visit window after the target day of Day 253). This upper analysis visit window incorporates the protocol-specified 5-day visit window for the PCP Week 36 visit, the additional protocol-specified 14-day allotment to allow for any delay in endoscopy assessment due to a pandemic (eg, Coronavirus Disease 2019 pandemic) or other global health emergency, and the protocol-specified requirement that the PCP Week 36 visit be conducted in-clinic within 3 days post-endoscopy.

If a subject's last assessment is after the date of first dose of PCP study treatment but prior to the first scheduled post-baseline visit/assessment, that assessment will be assigned to the first post-baseline scheduled visit. If two or more assessments occur in the same visit window, the assessment closest to the target visit day will be selected for inclusion in the analysis. If multiple assessments are equally close to the target visit day, then the latest assessment will be selected for inclusion in the analysis. If multiple assessments occur on the same day, the average of these assessments will be used for analysis (for evaluations with continuous values), excluding

analyses evaluated at any time post-baseline (eg, maximum post-baseline, shift from baseline, abnormality, and outlier analyses).

Further details regarding analysis visit windowing for efficacy endpoints are described in Section 8.5.1.

## 3.3. Standard Calculations

Standard calculations are described as follows:

Age:

Either date of birth or age at informed consent (as an integer) is collected in the IRT system for a given subject. If date of birth is collected, age at time of first dose of PCP study treatment will be calculated in years as [(date of first dose of PCP study treatment – date of birth +1)]/365.25, rounded down to the nearest integer. If age at informed consent is collected, age at time of first dose of PCP study treatment will be set equal to the reported age at informed consent.

Change/Percent Change from Baseline:

Change from baseline will be calculated as: Value at post-baseline visit – value at baseline.

Percent change from baseline will be calculated as: (Change from baseline/value at baseline) * 100%.

If the value at baseline is 0, percent change from baseline will be missing.

Study Day:

Study day will be calculated as follows:

Date of assessment – Date of first dose of PCP study treatment, where date of assessment is prior to the date of first dose of PCP study treatment

Date of assessment – Date of first dose of PCP study treatment + 1, where date of assessment is on or after the date of first dose of PCP study treatment.

Duration:

Where needed, the duration between two dates Date1 and Date2 will be calculated as follows:

Date2 – Date1 + 1, when expressed in days

(Date2 – Date1 + 1)/7, when expressed in weeks.

## 3.4. Considerations Related to the COVID-19 Pandemic

Screening of subjects for this study began in September 2020. As such, this study's conduct may be affected by the global Coronavirus Disease 2019 (COVID-19) pandemic. Prospectively specified measures to allow for adjustment in study conduct due to the COVID-19 pandemic (eg, measures addressing missed visits, alternative methods of performing study procedures, and alternative methods of supplying IP to subjects), with the objectives of maximizing subject monitoring, data collection, and endpoint ascertainment, are described in Protocol Appendix 9.

Several regulatory authorities have issued guidances related to clinical trial conduct and associated methodological issues related to the effect of the COVID-19 pandemic, including, but not limited to, the Food and Drug Administration (FDA) (FDA, 2020a, Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic and FDA, 2020b, Statistical Considerations for Clinical Trials During the COVID-19 Public Health Emergency) and European Medicines Agency (EMA) (EMA, 2021, Guidance on the Management of Clinical Trials during the COVID-19 (Coronavirus) pandemic, Version 4 and EMA, 2020, Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials).

Potential effects on conduct of this study related to the COVID-19 pandemic include, but are not limited to, the following: subject discontinuation of study treatment and/or withdrawal from study; temporary site closure; inability or reduced ability for subjects to continue on IP due to disruptions in IP dispensing; extending protocol-defined windows for visits to enable endpoint ascertainment, which is addressed in Protocol Appendix 9; missed visits or missed assessments; and alternative ascertainment methods (eg, remote ascertainment via a telephone visit or virtual visit) used in visits or assessments for collection of data contributing to efficacy and/or safety endpoints.

A determination will be made at the end of study if additional statistical analyses are warranted, should conduct of the study have been affected by the COVID-19 pandemic to a substantial degree. This determination may be made by assessing various aspects, such as the number of subjects with missed visits or assessments for reasons related to the COVID-19 pandemic, the number of subjects with alternative ascertainment methods used in visits or assessments, the number of subjects who discontinued study treatment and/or withdrew from the study for reasons related to the COVID-19 pandemic, and the number of subjects for whom IP dispensing was affected (eg, interruption in IP dispensing or alternative methods for supplying IP) for reasons related to the COVID-19 pandemic. Summary tables and/or by-subject data listings may be provided as part of this assessment.

If the determination of this assessment is that conduct of the study was substantially affected by the COVID-19 pandemic, additional statistical analyses may be undertaken to enable an understanding of the potential impact of the COVID-19 pandemic on study results, particularly in the scenario of statistically significant results for the primary and/or secondary endpoints.

A non-exhaustive list of examples of such additional statistical analyses is as follows:

- Analysis of compliance:
  - By subgroup of subjects for whom IP dispensing or accountability was affected for reasons related to the COVID-19 pandemic versus subjects for whom IP dispensing was not affected;
- Analyses of primary and select secondary efficacy endpoints:
  - By subgroup of subjects whose study participation was affected by the COVID-19 pandemic versus subjects whose study participation was not affected;
  - Excluding only data that is missing for reasons related to the COVID-19 pandemic (but including imputation of data missing for other reasons); and

  - By subgroup of subjects who had an alternative method of collection for an assessment contributing to a particular endpoint (eg, mobile endoscopy) versus subjects who did not have an alternative method of collection

- Analysis of AEs:
  - Incidence of AEs by subgroup of subjects whose study participation was affected by the COVID-19 pandemic versus subjects whose study participation was not affected;
  - Incidence of AEs that were collected via an alternative method versus incidence of AEs that were not collected via an alternative method
- Analysis of laboratory values:
  - By subgroup of subjects whose study participation was affected by the COVID-19 pandemic versus subjects whose study participation was not affected; and
  - Excluding values resulting from local collection due to reasons related to the COVID-19 pandemic.

# 4. ANALYSIS POPULATIONS

## 4.1. All Enrolled Population

The all enrolled population will include all subjects with a non-missing date of informed consent. The all enrolled population will be utilized for analysis of subject disposition.

## 4.2. PCP Intent-to-Treat Population

The PCP intent-to-treat (ITT) Population will include all subjects who are randomized and receive at least 1 dose of PCP study treatment. Subjects will be grouped according to their randomized treatment, regardless of the treatment actually taken or received.

The PCP ITT population will be used for analyses of major protocol deviations, demographic and other baseline characteristics, UC treatment history, concomitant UC background medications at baseline, and efficacy.

## 4.3. PCP Safety Population

The PCP safety population will include all subjects who receive at least 1 dose of PCP study treatment, with subjects grouped according to their actual treatment.

Determination of actual treatment will be as follows. Subjects who are randomized to placebo or GB004 480 mg QD will be grouped according to their randomized treatment group, except for cases where a subject receives and takes IP from one or more misdispensed IP bottle(s) (defined as an IP bottle which was not assigned to the subject by the IRT system) whose treatment group identity differs from randomized treatment. In such cases, subjects randomized to placebo will be grouped in the GB004 480 mg QD or GB004 480 mg BID treatment groups only if GB004 was actually taken, with treatment grouping determined according to the greater frequency (QD or BID) of GB004 actually taken based on study treatment accountability data, and subjects

randomized to GB004 480 mg QD will be grouped in the GB004 480 mg BID treatment group only if GB004 was actually taken BID based on study treatment accountability data.

The PCP safety population will be utilized for analyses of medical history, prior and concomitant medications, duration of study treatment, compliance, prescribed dose modifications, time on study, and safety.

# 5. STUDY SUBJECTS

## 5.1. Disposition of Subjects

The disposition of subjects will be summarized. Disposition and screen failure data will also be presented in by-subject data listings. In addition, a listing of the randomization scheme will be presented.

### 5.1.1. Placebo-Controlled Period Week 12 Analysis

The following counts and percentages of subjects will be summarized:

- Screened
- Screen failures, including screen failure reasons
- Randomized
- Randomized in Error, defined as randomized subjects who have any inclusion/exclusion criteria not met, including those specific inclusion/exclusion criteria not met
- Treated
- Analysis populations (All Enrolled, PCP ITT, and PCP Safety)
- Protocol version of informed consent (Version 1.0 or Version 2.0)
- Completed PCP Week 12 on study treatment
  - Met protocol-defined UC disease activity criteria at PCP Week 12
- Did not complete PCP Week 12 on study treatment, including reasons for study treatment discontinuation prior to PCP Week 12
- Completed PCP Week 12
- Did not complete PCP Week 12, including reasons for withdrawal prior to PCP Week 12

If a subject who consented to version 2.0 of the protocol did not complete study treatment in the PCP, a determination of whether the subject completed PCP Week 12 on study treatment will be made based on the subject's date of last dose of study treatment in the PCP relative to the subject's nominal PCP Week 12 visit date or relative to the target day of Day 85 for the PCP Week 12 visit if the nominal PCP Week 12 visit date is missing.

Similarly, if a subject who consented to version 2.0 of the protocol did not complete the PCP, a determination of whether the subject completed PCP Week 12 will be made based on the subject's last assessment date relative to the subject's nominal PCP Week 12 visit date or relative to the target day of Day 85 for the PCP Week 12 visit if the nominal PCP Week 12 visit date is missing.

### 5.1.2. Placebo-Controlled Period Final Analysis

In addition to the above described counts and percentages for the PCP Week 12 Analysis, the following counts and percentages of subjects will also be presented for the PCP Final Analysis:

- Completed PCP study treatment
- Did not complete PCP study treatment, including reasons for PCP study treatment discontinuation
- Completed PCP
- Did not complete PCP, including reasons for withdrawal from PCP

## 5.2. Major Protocol Deviations

Protocol deviations will be identified and reviewed on an ongoing basis by the study team and entered into a Clinical Trial Management System. Major protocol deviations are defined as those that can affect efficacy and/or safety assessments, the safety or mental integrity of a subject, or the scientific value of the study. Protocol deviations will be prospectively classified as major or minor.

The number and percentage of subjects with a major protocol deviation overall and for each major protocol deviation type will be summarized. In addition, COVID-19 related and non-COVID-19 related major protocol deviations will be summarized.

Protocol deviations will also be presented in a by-subject data listing.

In addition, subjects who were randomized in error and subjects who were mis-stratified (defined as subjects for whom the value of the randomization stratification factor, systemic corticosteroid use at baseline [yes/no], is different than the actual value of the stratification factor at randomization) will be presented in by-subject data listings.

# 6. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

## 6.1. Demographic and Baseline Characteristics

The following demographic and baseline characteristics will be summarized:

- Age (years), both continuously and categorically ($\geq 18$ - $< 50$, $\geq 50$ - $< 65$, $\geq 65$)
- Sex (male, female)
- Race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, Other, White)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino)

- Weight (kg)
- Height (cm)
- Body mass index (BMI) (kg/$m^2$), both continuously and categorically (< 18.5, ≥ 18.5 - < 25, ≥ 25 - < 30, ≥ 30)
- Region/Country (Australia, Europe [Czech Republic, Georgia, Italy, Moldova, Poland, Romania, Russia, Serbia, Ukraine], South Korea, and United States)

Demographic and baseline characteristics will also be presented in a by-subject data listing.

## 6.2. Baseline Disease Characteristics

The following baseline disease characteristics will be summarized:

- Extent of disease (left side of colon, extensive)
- Time since UC symptom onset (years), both continuously and categorically, using categories of < 1, ≥ 1 - < 3, ≥ 3 - < 7, and ≥ 7
- Time since UC diagnosis (years), both continuously and categorically, using the same categories as time since UC symptom onset
- Systemic corticosteroid use (yes, no)
  - Per randomization stratum
  - Actual value
- Systemic corticosteroid dependent
- Smoking history (current , former, never)
- Fecal calprotectin (mg/kg), both continuously and categorically (normal [< 50], high [≥ 50]; ≤ 250, > 250; and ≤ 500, > 500)
- Fecal lactoferrin (μg/g), both continuously and categorically (normal [≤ 7.24], high [> 7.24])
- C-reactive protein (mg/L), both continuously and categorically (normal [≤ 5.0], high [>5.0])
- UC-100 index

Baseline disease characteristics will also be presented in a by-subject data listing.

## 6.3. Baseline Mayo Score and Subscores

Baseline Mayo score and subscores will be summarized based on the following:

- Mayo score, both continuously and categorically
  - 0 - 2 (Quiescent disease), including individual categories of 0, 1, and 2
  - 3 - 5 (Mild disease), including individual categories of 3, 4, and 5
  - 6 - 10 (Moderate disease), including individual categories of 6, 7, 8, 9, and 10

  - 11 - 12 (Severe disease), including individual categories of 11 and 12

- Mayo subscores, both continuously and categorically
  - Stool frequency subscore (SFS)
    - 0 (Normal number of stools)
    - 1 (1 - 2 stools more than normal)
    - 2 (3 - 4 stools more than normal)
    - 3 (5 or more stools than normal)
  - Rectal bleeding subscore (RBS)
    - 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)
  - Endoscopic subscore
    - 0 (Normal or inactive disease)
    - 1 (Mild disease [erythema, decreased vascular pattern])
    - 2 (Moderate disease [marked erythema, absent vascular pattern, friability, erosions])
    - 3 (Severe disease [spontaneous bleeding, ulceration])
  - Physician's Global Assessment (PGA)
    - 0 (Normal)
    - 1 (Mild disease)
    - 2 (Moderate disease)
    - 3 (Severe disease)

- Modified Mayo score, both continuously and categorically
  - 0 - 1 (Quiescent disease), including individual categories of 0 and 1
  - 2 - 4 (Mild disease), including individual categories of 2, 3, and 4
  - 5 - 7 (Moderate disease), including individual categories of 5, 6, and 7
  - 8 - 9 (Severe disease), including individual categories of 8 and 9

- Partial Mayo score, both continuously and categorically
  - 0 - 1 (Quiescent disease)

  - 2 - 4 (Mild disease)
  - 5 - 7 (Moderate disease)
  - 8 - 9 (Severe disease)

- 6-Point Mayo score, both continuously and categorically
  - 0 (Quiescent disease)
  - 1 - 2 (Mild disease)
  - 3 - 4 (Moderate disease)
  - 5 - 6 (Severe disease)

- Reference normal stool frequency (number of stools per day), continuously and categorically
  - When in remission
  - Prior to UC symptom onset/diagnosis

Baseline Mayo scores and subscores will also be presented in a by-subject data listing.

## 6.4. Baseline Histologic Activity

Baseline histologic activity will be summarized based on the following:

- RHI, both continuously and categorically
  - 0 - 3 (Quiescent)
  - 4 - 10 (Mild)
  - 11 - 18 (Moderate)
  - 19 - 33 (Severe)

- RHI subscores, both continuously and categorically
  - Chronic inflammatory infiltrate RHI subscore
    - 0 (No increase)
    - 1 (Mild but unequivocal increase)
    - 2 (Moderate increase)
    - 3 (Marked increase)
  - Lamina propria neutrophils RHI subscore
    - 0 (None)
    - 1 (Mild but unequivocal increase)
    - 2 (Moderate increase)

  - 3 (Marked increase)
- Neutrophils in epithelium RHI subscore
  - 0 (None)
  - 1 (< 5% crypts involved)
  - 2 (< 50% crypts involved)
  - 3 (> 50% crypts involved)
- Erosion or ulceration RHI subscore
  - 0 (No erosion, ulceration or granulation tissue)
  - 1 (Recovering epithelium + adjacent inflammation; Probable Erosion - focally stripped)
  - 2 (Unequivocal erosion)
  - 3 (Ulcer or granulation tissue)

- Lamina propria neutrophils RHI subscore and neutrophils in epithelium RHI subscore, categorically (> 0 and > 0, > 0 and 0, 0 and > 0, 0 and 0, respectively)

- Geboes score
  - 0 (architectural changes)
  - 1 (chronic inflammatory infiltrate)
  - 2 (lamina propria neutrophils)
  - 3 (neutrophils in epithelium)
  - 4 (crypt destruction)
  - 5 (erosions or ulcerations)

Baseline histologic activity will also be presented in a by-subject data listing.

## 6.5. UC Treatment History

UC treatment history will be summarized as the number and percentage of subjects with any prior use of the following medication types since UC diagnosis, including information on specific medications taken for each medication type:

- Aminosalicylates;
- Corticosteroids; and
- Immunomodulators.

The number and percentage of subjects with prior failure to respond and/or prior intolerance, with prior failure to respond, and with prior intolerance will be summarized for each of the above medication types.

Ulcerative colitis treatment history will also be presented in a by-subject data listing.

## 6.6. Medical History

Medical history (including surgical history) will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) version 24.0 or later. Medical history will be summarized by SOC and PT, with sorting by the internationally agreed order for SOC, and by decreasing frequency and then alphabetically for PT within SOC in the overall column. Counting will be at the subject level for each level of summarization (eg, any medical history, SOC, and PT), with subjects experiencing more than one occurrence of a SOC or PT counted only once.

Medical history will also be presented in a by-subject data listing.

## 6.7. Prior and Concomitant Medications

Prior medications are defined as all medications that started prior to the date of first dose of PCP study treatment. Concomitant medications are defined as all medications that started on or after the date of first dose of PCP study treatment or that started prior to the date of first dose of PCP study treatment and stopped on or after the date of first dose of PCP study treatment or are ongoing at any time during the PCP. A medication may be considered as both prior and concomitant (ie, prior and concomitant medications are not mutually exclusive).

In general, if it is not clear whether a medication is prior and/or concomitant due to missing or incomplete medication start and/or stop dates, the medication will be considered to be prior and/or concomitant for the given analysis, unless the non-missing portions of the medication start and/or stop dates indicate otherwise.

For subjects who consent to version 2.0 of the protocol, in cases where a medication has a non-missing start date equal to the date of PCP Week 12, the medication will be considered concomitant for the PCP Week 12 Analysis if it is reported on the appropriate concomitant medication eCRF intended to capture medications through PCP Week 12. If the date of PCP Week 12 is missing, then the target day for PCP Week 12 (Day 85) will be utilized for the purpose of determining concomitancy.

In cases where a medication has a non-missing start date equal to the date of first dose of OLE study treatment, the medication will be considered concomitant to the PCP if it is reported on the appropriate concomitant medication eCRF(s) intended to capture medications in the PCP.

Verbatim prior and concomitant medication terms will be coded to a drug class (anatomical therapeutic chemical classification level 2 [ATC2]) and preferred name using World Health Organization (WHO) Drug Global B3 March 2021 or later.

Separate summaries will be presented for prior medications and concomitant medications. Counting will be at the subject level for each level of summarization (eg, any medication, drug class, and preferred name), with subjects receiving more than one medication counted only once.

Prior and concomitant medications will also be presented in a by-subject data listing and will include an indicator, identifying each medication as prior and/or concomitant.

Concurrent procedures such as colectomy and ostomy will be presented in a separate by-subject data listing.

### 6.7.1. Concomitant Background UC Medications

Concomitant background UC medications at baseline will be summarized as the number and percentage of subjects with concomitant use of aminosalicylates and concomitant use of systemic corticosteroids at baseline, including specific medications taken for each medication type.

The number and percent of subjects who are steroid-free at PCP Week 36, defined as not requiring any treatment with a systemic corticosteroid for a minimum duration prior to PCP Week 36, will be evaluated in the subset of subjects in the ITT population with an actual stratification value of yes for systemic corticosteroid use at baseline. Two different minimum durations will be evaluated, 2 weeks and 4 weeks. Steroid-free duration prior to PCP Week 36 may be summarized continuously as well.

# 7. EXTENT OF EXPOSURE, COMPLIANCE, AND TIME ON STUDY

## 7.1. Duration of Study Treatment

Duration of PCP study treatment will be summarized, both as a continuous parameter (in units of days and weeks, as well as total number of subject-years) and as a categorical parameter as follows:

- PCP Week 12 Analysis
  - $< 2$ weeks, $< 4$ weeks, $\geq 2$ weeks, $\geq 4$ weeks, $\geq 8$ weeks, $\geq 11$ weeks, 2 days (which is the beginning of the protocol-specified visit window for the PCP Week 12 visit), and $\geq 12$ weeks
- PCP Final Analysis
  - The above categories for the PCP Week 12 Analysis, plus $\geq 18$ weeks, $\geq 24$ weeks, $\geq 35$ weeks, 2 days (which is the beginning of the protocol-specified visit window for the PCP Week 36 visit), and $\geq 36$ weeks

Other categories of duration of PCP study treatment may be utilized for summary purposes as needed.

Duration of PCP study treatment in days will be calculated for the PCP Week 12 Analysis as follows:

- For subjects who consent only to protocol version 1.0 and for subjects who consent to protocol version 2.0 and do not complete PCP Week 12 on study treatment:
  - If the subject's time of last dose of PCP study treatment is prior to 12:00 noon: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 0.5
  - If the subject's time of last dose of PCP study treatment is after 12:00 noon or is missing: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 1.
- For subjects who complete PCP Week 12 on study treatment:

- If the subject received in clinic dosing at the PCP Week 12 nominal visit: Date of PCP Week 12 nominal visit – Date of first dose of PCP study treatment + 0.5 (due to the last dose of PCP study treatment as of the PCP Week 12 visit being administered per protocol as a morning dose in clinic at the PCP Week 12 nominal visit)
- If the subject had a PCP Week 12 nominal visit but did not receive in clinic dosing at the PCP Week 12 nominal visit:
  - If the subject did not continue dosing in the PCP beyond the PCP Week 12 nominal visit:
    - If the subject's time of last dose of PCP study treatment is prior to 12:00 noon: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 0.5
    - If the subject's time of last dose of PCP study treatment is after 12:00 noon or is missing: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 1
  - If the subject continued dosing in the PCP beyond the PCP Week 12 nominal visit:
    - Date of PCP Week 12 nominal visit – Date of first dose of PCP study treatment
- If the subject has a missing date for the PCP Week 12 nominal visit, then duration of PCP study treatment will be imputed to 84.5.

Duration of PCP study treatment in days will be calculated for the PCP Final Analysis as follows:

- If the subject entered OLE and the subject's date of last dose of PCP study treatment is the same as the subject's first dose of OLE study treatment: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 0.5
- If the subject entered OLE and the subject's date of last dose of PCP study treatment is not the same as the subject's first dose of OLE study treatment, or if the subject did not enter OLE:
  - If the subject's time of last dose of PCP study treatment is prior to 12:00 noon: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 0.5
  - If the subject's time of last dose of PCP study treatment is on or after 12:00 noon or is missing: Date of last dose of PCP study treatment – Date of first dose of PCP study treatment + 1

Duration of study treatment will also be presented in a by-subject data listing.

## 7.2. Compliance

Compliance will be assessed using the following formula:

Compliance (%) = (Number of actual tablets taken/Number of expected tablets) x 100,

where:

Number of actual tablets taken = (Number of tablets dispensed – Number of tablets returned); and Number of expected tablets is based on actual duration of study treatment as described in Section 7.1.

Compliance will be calculated using two different parameters, for which the determination of the number of expected tablets differs. The first parameter, termed compliance with prescribed PCP study treatment, will be calculated taking into account prescribed dose modifications to calculate the number of expected tablets; that is, the number of prescribed tablets will be used as the value for the number of expected tablets. This first parameter represents the extent to which a subject complies with prescribed PCP study treatment. The second parameter, termed compliance with randomized PCP study treatment, will be calculated without taking into account prescribed dose modifications to calculate the number of expected tablets. This second parameter therefore represents the extent to which a subject adhered to randomized PCP study treatment.

Compliance will be summarized using descriptive statistics and categorically (eg, < 80%, ≥ 80 - 100%, > 100 - < 120%, ≥ 120%). Further analyses of compliance may be performed according to compliance with tablets versus delayed-release tablets.

Compliance will also be presented in a by-subject data listing. In addition, IP accountability including bottle information will be presented in a by-subject data listing.

## 7.3. Prescribed Dose Modifications

Provisions for permitted prescribed dose modifications of IP are described in Protocol Section 6.5.1.

The number and percentage of subjects with a prescribed dose modification, defined as the occurrence of a prescribed dose reduction or prescribed dose interruption, will be summarized.

Subjects with occurrence of a prescribed dose modification will be summarized according to:

- Reasons for prescribed dose modification (liver chemistry abnormality, treatment-related AE other than liver chemistry abnormality, and other).
- Prescribed dose modification type:
  - Those who had only a first step prescribed dose reduction (defined as a reduction in IP dose by 2 tablets BID to a total of 6 tablets BID), and those who had a second step prescribed dose reduction (defined as a further reduction in IP dose by 1 tablet BID to a total of 5 tablets BID).
  - Those who had a prescribed dose interruption.
- Prescribed dose resumption following a prescribed dose modification:
  - Those who resumed a higher dose.

  - Those who resumed the original dose.

- Completion of PCP study treatment following a prescribed dose modification:
  - Those who completed PCP study treatment following a prescribed dose modification, broken down by those who completed at a reduced dose versus those who completed at the original dose.
  - Those who discontinued PCP study treatment following a prescribed dose modification.

  For the PCP Week 12 Analysis, this will be summarized with respect to completion of PCP Week 12 on study treatment, and for the PCP Final Analysis, this will be summarized with respect to completion of PCP Week 12 on study treatment and completion of PCP on study treatment.
- Duration of prescribed dose modification, defined as time in days from the start of a prescribed dose modification to resumption of original dose (for subjects who resume original dose) or to last dose of study treatment (for subjects who do not resume original dose at time of last dose). For subjects with multiple occurrences of prescribed dose modification, the total duration of prescribed dose modification will be summarized.

Prescribed dose modification information will also be presented in a by-subject data listing.

## 7.4. Time on Study

Duration of time on study in the PCP will be summarized, both as a continuous parameter (in units of days and weeks, as well as total number of subject-years) and as a categorical parameter as follows:

- PCP Week 12 Analysis
  - $< 2$ weeks, $< 4$ weeks, $\geq 2$ weeks, $\geq 4$ weeks, $\geq 8$ weeks, $\geq 11$ weeks, 2 days (which is the beginning of the protocol-specified visit window for the PCP Week 12 visit), $\geq 12$ weeks, and $\geq 16$ weeks
- PCP Final Analysis
  - The above categories for the PCP Week 12 Analysis through $\geq 12$ weeks inclusive, plus $\geq 18$ weeks, $\geq 24$ weeks, $\geq 35$ weeks, 2 days (which is the beginning of the protocol-specified visit window for the PCP Week 36 visit), $\geq 36$ weeks, $\geq 39$ weeks, 4 days (which is the beginning of the protocol-specified visit window for the PCP Week 40 Follow-up visit), and $\geq 40$ weeks

Other categories of duration of PCP time on study may be utilized for summary purposes as needed.

Duration of PCP time on study in days will be calculated as:

- PCP Week 12 Analysis:
  - For subjects who consent only to protocol version 1.0 and for subjects who consent to protocol version 2.0 and withdraw from PCP at or prior to PCP

Week 12: Date of last assessment in PCP – Date of first dose of PCP study treatment + 1

- For subjects who consent to protocol version 2.0 and do not withdraw from PCP at or prior to PCP Week 12: Date of PCP Week 12 nominal visit – Date of first dose of PCP study treatment + 1; if the date of PCP Week 12 nominal visit is missing, then time on study will be imputed to 85.

- PCP Final Analysis:
  - For subjects who are dosed in the OLE: Date of first dose of OLE study treatment – Date of first dose of PCP study treatment + 1
  - For subjects who are not dosed in the OLE: Date of last assessment in PCP – Date of first dose of PCP study treatment + 1.

Duration of PCP time on study will also be presented in a by-subject data listing.

# 8. EFFICACY EVALUATION

## 8.1. Overview of Efficacy Analysis Issues

### 8.1.1. Handling of Dropouts or Missing Data

The below approaches for handling of missing data will be applied after employing the approach for handling intercurrent events (ICEs) (ie, treatment failure rules) described in Section 8.1.4 below.

#### 8.1.1.1. Proportion-Based (Binary) Efficacy Endpoints

The primary analysis approach for handling of missing data for proportion-based (binary) efficacy endpoints will be non-responder imputation (NRI), wherein subjects with a missing value due to insufficient data for determination of endpoint status, including missing baseline value(s) and/or post-baseline value(s), will be considered as having not met the endpoint.

#### 8.1.1.2. Change from Baseline (Continuous) Efficacy Endpoints

As all change from baseline (continuous) efficacy endpoints are exploratory endpoints, primary analyses of change from baseline efficacy endpoints will not employ imputation for missing data. Additional analyses of change from baseline efficacy endpoints (eg, mixed-effects model with repeated measures, methods utilizing imputation for missing data) may be explored.

### 8.1.2. Multicenter Studies

Data from all countries and sites will be pooled for the purpose of analyses.

### 8.1.3. Multiple Comparisons/Multiplicity

In general, statistical hypothesis testing for efficacy endpoints will be made without any adjustments for multiplicity, and all statistical hypothesis tests will be performed using a two-sided significance level of 0.050.

### 8.1.4. Handling of Intercurrent Events (Treatment Failure Rules)

For efficacy analyses, handling of intercurrent events (also known as treatment failure rules) will be applied as described below.

Subjects with any of the following ICEs that occur prior to the timepoint of the endpoint will be considered as not meeting the endpoint for proportion-based (binary) endpoints and as having missing values for change from baseline-based endpoints:

- Protocol-prohibited post-baseline initiation of a new UC rescue therapy medication
  - This includes post-baseline initiation of new UC rescue therapy medications such as systemic corticosteroids, azathioprine, 6-mercaptopurine, biologic therapies including anti-tumor necrosis factor agents or vedolizumab, tofacitinib, oral cyclosporine, sirolimus, mycophenolate mofetil, and per rectum therapy (eg, enemas or suppository formulations of steroids or 5-aminosalicylates) except those required for flexible sigmoidoscopy preparation. Use of systemic corticosteroids and per rectum therapy initiated post-baseline must be for > 3 consecutive days in order to qualify for this ICE.
  - This ICE is considered to have occurred on the first start date of new UC rescue therapy medication use.
- Protocol-prohibited post-baseline increase in a concomitant background UC therapy used at baseline (as described in Protocol Section 6.5.2)
  - This includes the following:
    - Post-baseline increase of ≥ 20% in the average total daily dose of concomitant background aminosalicylate UC treatment for ≥ 7 consecutive days, relative to the maximum total daily dose from 2 weeks prior to screening endoscopy up through and including the day prior to first dose of PCP study treatment (Day -1).
    - Post-baseline increase in the total daily dose of concomitant background systemic corticosteroid UC treatment for > 3 consecutive days, relative to the maximum total daily dose from 2 weeks prior to screening endoscopy up through and including the day prior to first dose of PCP study treatment (Day -1). Protocol-permitted increases in systemic corticosteroid dose up to baseline dose for a subject on systemic corticosteroids at baseline in whom systemic corticosteroid taper is initiated after PCP Week 12 and who cannot tolerate the systemic corticosteroid taper without recurrence of clinical symptoms of either UC or steroid withdrawal will not be considered an ICE.
  - This ICE is considered to have occurred on the first date of qualifying dose increase.
- Post-baseline prolonged course of systemic corticosteroids, where ‘prolonged’ is defined as consecutive use for > 7 days, for an indication other than UC.

  This ICE is considered to have occurred on the first start date of qualifying systemic corticosteroid use.

- A colectomy (partial or total) or an ostomy.

  This ICE is considered to have occurred on the procedure date for the colectomy or ostomy.

- Discontinuation of PCP study treatment > 28 days prior to the timepoint of the endpoint.

The approach for handling intercurrent events described above will be applied prior to employing the approaches for handling of missing data described in Section 8.1.1 above.

Identification of ICEs may be performed by programmatic derivation and/or manual review. For the PCP Week 12 Analysis, this will be performed prior to Sponsor unblinding.

The number and percentage of subjects with an ICE will be summarized overall and by each of the types of ICEs listed above. For the PCP Final Analysis, this will be further summarized with respect to ICE occurrence prior to PCP Week 12. A by-subject data listing of ICEs will be provided.

## 8.2. Efficacy Endpoints

Table 3 below provides a summary of the statistical models/methods of primary and sensitivity analyses to be used for primary, secondary, and exploratory efficacy endpoints.

**Table 3: Efficacy Endpoints and Analysis Methods**

| Efficacy Endpoints | CMH | Logistic Regression | ANCOVA |
|---|---|---|---|
| Primary and Secondary | Primary Analyses, Sensitivity Analyses | Sensitivity Analyses | |
| Exploratory | | | |
| Proportion-based | Primary Analyses | | |
| Change from baseline | | | Primary Analyses |

Abbreviations: ANCOVA, analysis of covariance; CMHI, Cochran-Mantel Haenszel

Analyses of efficacy endpoints for the PCP Week 12 Analysis (ie, efficacy endpoints evaluated at or up to and including PCP Week 12) will be based on the ITT population, unless otherwise stated below.

Analyses of efficacy endpoints for the PCP Final Analysis involving timepoints after PCP Week 12 will be based on the subset of subjects in the ITT population who consent to version 2.0 of the protocol, unless otherwise stated below.

Further analysis population details for specific efficacy endpoints are described below.

Certain secondary and exploratory efficacy endpoints will only be evaluated in a subset of subjects in the ITT population in order to ensure that achievement of the endpoint can be attributed to study treatment and that treatment effect estimation is not affected by any potential imbalance across treatment groups with respect to subjects fully or partially meeting requirements for the endpoint at baseline. This includes the following:

- Endpoints assessing histologic remission (including mucosal healing and disease clearance) will be evaluated in the subset of subjects in the ITT population with both baseline RHI lamina propria neutrophils and neutrophils in the epithelium subscores > 0.
- Endpoints assessing histologic response will be evaluated in the subset of subjects in the ITT population with baseline RHI score ≥ 7.
- Endpoints assessing subjects with a Geboes score ≤ 2 will be evaluated in the subset of subjects in the ITT population with a baseline Geboes score > 2.
- The endpoint assessing steroid-free clinical remission at PCP Week 36 will be evaluated in the subset of subjects in the ITT population with an actual stratification value of yes for systemic corticosteroid use at baseline.
- Endpoints assessing the proportion of subjects with RHI score ≤ 3 will be evaluated in the subset of subjects in the ITT population with baseline RHI score > 3.
- Endpoints assessing the proportion of subjects with fecal calprotectin, fecal lactoferrin, or CRP values below a certain threshold value will be evaluated in the subset of subjects in the ITT population who baseline value is above the threshold value for the endpoint.

For all applicable efficacy analyses, stratification factor values (systemic corticosteroid use at baseline [yes/no]) will be based on the values used for randomization, unless otherwise stated.

## 8.3. Analysis Methods

### 8.3.1. Primary Endpoint

#### 8.3.1.1. Primary Analysis

The primary objective of the PCP will be evaluated by testing the superiority of each GB004 treatment group to the placebo treatment group with respect to the primary endpoint of the proportion of subjects with clinical remission at PCP Week 12 based on a Cochran-Mantel Haenszel (CMH) chi-squared test (Mantel, 1959) stratified by the randomization stratification factor of systemic corticosteroid use at baseline (yes/no) with a two-sided 0.050 level of significance for each comparison.

The Mantel-Fleiss criterion (Mantel, 1980) will be used to assess the validity of the chi-square approximation for the distribution of the CMH test statistic. In the case that the Mantel-Fleiss criterion is not at least 5, indicating lack of validity of the chi-square approximation due to stratum cell sizes that are too small, Pearson's chi-squared test will be used instead of the CMH chi-squared test in the primary analysis of the primary endpoint. In the rare case that expected cell frequencies are too small, Fisher's exact test will be used instead of Pearson's chi-squared test.

For the comparison of each GB004 treatment group versus placebo, the p-value from the CMH chi-squared test (or Pearson's chi-squared test and/or Fisher's exact test, as described above) will be reported. The number and proportion of subjects meeting the primary endpoint and corresponding 95% Wilson (Score) CIs (Wilson, 1927) will be summarized by treatment group.

The unstratified, absolute difference in proportion between each GB004 treatment group and placebo will be summarized, with the 95% CI for the unstratified, absolute difference based on the Newcombe continuity-corrected method (Newcombe, 1998). The estimated Mantel-Haenszel common odds ratio across strata and corresponding 95% CI will also be summarized.

The primary analysis of the primary endpoint will be based on application of ICE handling followed by NRI.

A bar graph of the proportion of subjects meeting the primary endpoint by treatment group will be presented.

Primary endpoint data will be presented in a by-subject data listing.

#### 8.3.1.2. Sensitivity Analyses

Sensitivity analyses of the primary endpoint will be conducted to evaluate the robustness of treatment effects observed in the primary analysis of the primary endpoint described in Section 8.3.1.1 above.

Table 4 below provides an overview of the primary and sensitivity analyses of the primary endpoint.

**Table 4: Summary of Primary and Sensitivity Analyses of the Primary Endpoint**

| Analysis | Description |
|---|---|
| Primary Analysis:<br>Analysis with both ICE Handling and NRI Applied | • Analysis based on application of both ICE Handling and NRI.<br>• P-value based on CMH chi-squared test.<br>• Estimated treatment effect based on unstratified, absolute difference in proportion, with 95% CI based on the Newcombe continuity-corrected method. |
| Sensitivity Analysis #1:<br>Analysis with NRI Applied without ICE Handling Applied | • Analysis based on application of NRI without application of ICE Handling.<br>• P-value based on CMH chi-squared test.<br>• Estimated treatment effect based on unstratified, absolute difference in proportion, with 95% CI based on the Newcombe continuity-corrected method. |
| Sensitivity Analysis #2:<br>Analysis with ICE Handling Applied without NRI Applied | • Analysis based on application of ICE Handling without application of NRI.<br>• P-value based on CMH chi-squared test.<br>• Estimated treatment effect based on unstratified, absolute difference in proportion, with 95% CI based on the Newcombe continuity-corrected method. |
| Sensitivity Analysis #3:<br>As-Observed Analysis without ICE Handling or NRI Applied | • Analysis using observed data, without application of ICE Handling or NRI.<br>• P-value based on CMH chi-squared test.<br>• Estimated treatment effect based on unstratified, absolute difference in proportion, with 95% CI based on the Newcombe continuity-corrected method. |
| Sensitivity Analysis #4:<br>Logistic Regression with ICE Handling and NRI Applied | • Analysis based on application of Treatment Failure Rules and NRI.<br>• Estimated treatment effect and p-value based on OR from logistic regression analysis. |

Abbreviations: CI, confidence interval; CMH, Cochran-Mantel Haenszel; ICE, intercurrent event; NRI, non-responder imputation; OR, odds ratio.

#### 8.3.1.2.1. Sensitivity Analysis #1: NRI Applied without ICE Handling Applied

In this sensitivity analysis, the primary endpoint will be analyzed similarly as in the primary analysis but with application of NRI without ICE handling.

#### 8.3.1.2.2. Sensitivity Analysis #2: ICE Handling without NRI Applied

In this sensitivity analysis, the primary endpoint will be analyzed similarly as in the primary analysis, but with ICE handling applied without NRI applied.

#### 8.3.1.2.3. Sensitivity Analysis #3: As-Observed Analysis without ICE Handling or NRI Applied

In this sensitivity analysis, the primary endpoint will be analyzed similarly as in the primary analysis but based on observed data without application of ICE handling or NRI.

#### 8.3.1.2.4. Sensitivity Analysis #4: Logistic Regression with ICE Handling and NRI Applied

In this sensitivity analysis, logistic regression modeling will be used to compare each GB004 treatment group with the placebo treatment group, after handling of ICEs and NRI are applied.

The logistic regression model will include covariates for treatment group, the randomization stratification factor of systemic corticosteroid use at baseline (yes/no), baseline Modified Mayo score, and region (Europe vs Rest of World).

The odds ratio (OR), corresponding asymptotic 95% CI, and p-value from the logistic regression model will be presented for the comparison of each GB004 treatment group versus placebo.

If the logistic regression model does not converge, covariates will be removed from the model, and the first reduced model that converges among reduced models with the below specified covariates will be utilized in the following order of priority:

- Treatment group, randomization stratification factor, and baseline Modified Mayo score
- Treatment group and randomization stratification factor
- Treatment group only

### 8.3.2. Secondary Endpoints

#### 8.3.2.1. Primary Analysis

The secondary efficacy objectives of the PCP will be evaluated by testing the superiority of each GB004 treatment group to the placebo treatment group with respect to the secondary efficacy endpoints described and defined in Section 2 with a two-sided 0.050 level of significance for each comparison. The primary analyses of these secondary efficacy endpoints will be performed using the same approach as the primary analysis of the primary endpoint.

Secondary efficacy endpoint data will be presented in a by-subject data listing.

#### 8.3.2.2. Sensitivity Analyses

Sensitivity analyses of secondary efficacy endpoints will be conducted to evaluate the robustness of treatment effects observed in the primary analyses of secondary efficacy endpoints described in Section 8.3.2.1 above. These sensitivity analyses will be performed using the same approaches as the sensitivity analyses of the primary endpoint, with the differences described below.

For Sensitivity Analysis #4 (Logistic Regression with ICE Handling and NRI Applied), covariates in the logistic regression models for secondary efficacy endpoints will be the same as for the primary endpoint with the following exceptions:

- histologic remission will utilize the covariate of baseline RHI score instead of baseline Modified Mayo score;

- endoscopic improvement will utilize the covariate of baseline endoscopic subscore instead of baseline Modified Mayo score;
- mucosal healing will utilize the covariates of baseline RHI score and baseline endoscopic subscore instead of baseline Modified Mayo score.

If any of the logistic regression models for sensitivity analyses of secondary efficacy endpoints do not converge, the approach utilized for selection of a reduced model that converges will be similar to that for the primary endpoint. For mucosal healing, baseline endoscopic subscore will be removed from the model prior to removal of baseline RHI score.

### 8.3.3. Potential Additional Sensitivity Analysis of Primary and Secondary Endpoints

An additional sensitivity analysis of the primary and/or secondary endpoints may be performed, employing a hybrid imputation approach combining NRI and multiple imputation (MI) (Chen, 2021), particularly in the scenario of substantial missingness assessed to be due to an underlying missing at random (MAR) missing data mechanism assumption. This hybrid imputation approach has the advantage of combining missing data handling techniques under different missing data mechanism assumptions, including missing at random (MAR) and missing not at random (MNAR) assumptions, in order to evaluate the impact of missingness due to different reasons.

The MAR assumption is that missingness does not depend on the actual missing values, but that the missingness can be completely explained by the observed data. The MNAR assumption is that missingness does depend on the actual missing values even after conditioning on the observed data.

The MNAR assumption will be made for missingness caused by withdrawals due to reasons such as AEs or lack of efficacy and for subjects with an ICE prior to the timepoint of the endpoint. Missingness for other reasons (e.g., missingness caused by withdrawals for other reasons such as withdrawal by subject or loss to follow-up, including missingness arising from force majeure events such as wars) will be assumed to be MAR.

This hybrid imputation approach will be performed using the following 3 steps:

- Step 1:

  Subjects with missing endpoint analysis values are categorized into MNAR and MAR based on the above.

- Step 2:

  Missing endpoint analysis values are imputed using the statistical techniques under MNAR and MAR, respectively.

  NRI is used to impute subjects with missing endpoint analysis values under MNAR.

  For subjects with missing endpoint analysis values under MAR, MI is used to impute each Mayo subscore contributing to the endpoint using a fully conditional specification ordinal logistic regression method. Covariates included in this method may include Mayo subscores from all previous visits (including baseline), treatment group, and the randomization stratification factor. Multiple imputation will be utilized to compute a

number of multiply imputed datasets (e.g., 50) for each Mayo subscore. The Modified Mayo score and endpoint analysis values are then derived in each of the multiply imputed datasets.

- Step 3:

  The multiply imputed datasets under MNAR and MAR from Step 2 are combined to obtain complete datasets with non-missing endpoint analysis values for all subjects in the analysis population. The primary analysis method is conducted within each multiply imputed dataset, and the analysis results from the multiply imputed datasets are combined using Rubin's rules to obtain estimates for relevant statistics, including differences in proportion and associated 95% CIs and CMH p-values. Since the CMH test statistic is non-normally distributed, the Wilson-Hilferty transformation will be used for normalization before applying Rubin's rules.

### 8.3.4. Exploratory Endpoints

#### 8.3.4.1. Exploratory Efficacy Endpoints

Exploratory efficacy endpoint data will be presented in a by-subject data listing.

##### 8.3.4.1.1. Proportion-Based (Binary) Exploratory Efficacy Endpoints

Proportion-based (binary) exploratory endpoints will be analyzed similarly to the primary analysis of the primary endpoint.

A subject will be considered to have met the exploratory efficacy endpoint of steroid-free clinical remission at PCP Week 36 if the subject achieves clinical remission at PCP Week 36 without requiring any treatment with a systemic corticosteroid for a minimum duration prior to PCP Week 36. Two different minimum durations will be evaluated, 2 weeks and 4 weeks. Since steroid-free clinical remission at PCP Week 36 will be evaluated in the subset of subjects in the ITT population with an actual stratification value of yes for systemic corticosteroid use at baseline, analysis of steroid-free clinical remission at PCP Week 36 will utilize Pearson's chi-squared test rather than CMH.

##### 8.3.4.1.2. Change from Baseline (Continuous) Exploratory Efficacy Endpoints

Change from baseline (continuous) exploratory endpoints will be analyzed using analysis of covariance (ANCOVA) models at all applicable timepoints based on application of ICE handling (where subjects with an ICE that occurs prior to the timepoint of the endpoint will be considered as having missing values as described in Section 8.1.4). All ANCOVA models will include covariates for treatment group, the randomization stratification factor of systemic corticosteroid use at baseline (yes/no), and baseline value corresponding to the change from baseline endpoint.

Descriptive statistics for change from baseline will be presented, along with the LS means, SEs, and corresponding 95% CIs for each treatment group from the ANCOVA models. The difference in LS means, SE, corresponding 95% CI, and p-value for each GB004 treatment group versus the placebo group from the ANCOVA models will also be presented.

#### 8.3.4.2. Pharmacokinetic Concentrations

By-subject data listings of pharmacokinetic plasma, tissue, and fecal concentrations for GB004 (and its metabolites, as applicable) will be presented.

## 8.4. Examination of Subgroups

Subgroup analyses of the primary and secondary efficacy endpoints will be performed as described below. These subgroup analyses are aimed at evaluation of consistency of treatment effects across subgroup levels, rather than statistical inference within subgroup levels.

Subgroup analyses will be performed similarly to the primary analysis methods for these endpoints with respect to application of ICE handling followed by NRI. Pearson's chi-squared test will be utilized to calculate within subgroup level p-values, unless expected cell frequencies are too small, in which case Fisher's exact test will be used. The Breslow-Day test with Tarone's adjustment (Tarone, 1985) will be used to evaluate the homogeneity of odds ratios across levels of each subgroup.

Subgroup analyses will summarize, for each subgroup level, the numbers and proportions of subjects meeting the endpoint by treatment group and corresponding 95% Wilson (Score) CIs, absolute differences in proportions between each GB004 treatment group and placebo and corresponding 95% CIs based on the Newcombe continuity-corrected method, odds ratios and corresponding 95% CIs, and within subgroup level p-values (as described above), as well as p-values from the Breslow-Day test with Tarone's adjustment. Subgroup analyses will be presented in tables and forest plots.

For any subgroup level that does not comprise ≥ 10% of the ITT population (or the applicable subset of the ITT population - eg, for histologic remission and mucosal healing), only within treatment group proportions and corresponding 95% Wilson (Score) CIs will be presented for that subgroup level, with no statistics for treatment effect comparisons within subgroup level. The Breslow-Day test with Tarone's adjustment will still be performed in such cases if there are at least 2 subgroup levels comprising ≥ 10% of the ITT population (or the applicable subset of the ITT population).

Subgroup analyses of interest are described in Table 5 below.

**Table 5: Subgroup Analyses**

| Subgroup Category | Subgroups | Efficacy Endpoints to be Evaluated |
|---|---|---|
| Demographic and baseline characteristics | • Age (years; < 50, ≥ 50)<br>• Sex (male, female)<br>• BMI (kg/m2; < 25, ≥ 25)<br>• Region (Europe, Rest of World) | • Proportion of subjects with clinical remission at PCP Week 12<br>• Proportion of subjects with clinical response at PCP Week 12<br>• Proportion of subjects with clinical remission at PCP Week 36<br>• Proportion of subjects with clinical response at PCP Week 36 |
| Baseline disease characteristics | • Extent of disease (left side of colon, extensive)<br>• Time since UC diagnosis (years; < 3, ≥ 3)<br>• Systemic corticosteroid use ([per randomization stratum and per actual value]; no, yes)<br>• Systemic corticosteroid dependent (no, yes)<br>• Prior use of immunomodulators (no, yes)<br>• Smoking history (never, former, current)<br>• Fecal calprotectin (mg/kg; ≤ Median, > Median, where median is based on the ITT population)<br>• Fecal lactoferrin (mg/kg; ≤ Median, > Median, where median is based on the ITT population)<br>• CRP (mg/L; ≤ Median, > Median, where median is based on the ITT population) | • Proportion of subjects with clinical remission at PCP Week 12<br>• Proportion of subjects with clinical response at PCP Week 12<br>• Proportion of subjects with clinical remission at PCP Week 36<br>• Proportion of subjects with clinical response at PCP Week 36 |
| Baseline Mayo and Modified Mayo score subgroups | • Baseline Mayo score<br>– 5, 6 - 10<br>– 5 - 7, 8 - 10<br>• Baseline Modified Mayo score<br>– 4, 5 - 8<br>– 4, 5 - 7, 8<br>– 4 - 5, 6, 7 - 8 | • Proportion of subjects with clinical remission at PCP Week 12<br>• Proportion of subjects with clinical response at PCP Week 12<br>• Proportion of subjects with clinical remission at PCP Week 36<br>• Proportion of subjects with clinical response at PCP Week 36 |

| Subgroup Category | Subgroups | Efficacy Endpoints to be Evaluated |
|---|---|---|
| Baseline endoscopic subscore subgroups | • Baseline endoscopic subscore (2, 3) | • Proportion of subjects with clinical remission at PCP Week 12<br>• Proportion of subjects with clinical response at PCP Week 12<br>• Proportion of subjects with histologic remission at PCP Week 12<br>• Proportion of subjects with endoscopic improvement at PCP Week 12<br>• Proportion of subjects with mucosal healing at PCP Week 12<br>• Proportion of subjects with clinical remission at PCP Week 36<br>• Proportion of subjects with clinical response at PCP Week 36<br>• Proportion of subjects with histologic remission at PCP Week 36<br>• Proportion of subjects with endoscopic improvement at PCP Week 36<br>• Proportion of subjects with mucosal healing at PCP Week 36 |
| Baseline histologic activity subgroups | • Baseline RHI score<br>– 0 - 3, > 3<br>– 0 - 10, > 10<br>• Both baseline RHI lamina propria neutrophils and neutrophils in the epithelium subscores > 0 (yes, no) | • Proportion of subjects with clinical remission at PCP Week 12<br>• Proportion of subjects with clinical response at PCP Week 12<br>• Proportion of subjects with histologic remission at PCP Week 12<br>• Proportion of subjects with endoscopic improvement at PCP Week 12<br>• Proportion of subjects with mucosal healing at PCP Week 12<br>• Proportion of subjects with clinical remission at PCP Week 36<br>• Proportion of subjects with clinical response at PCP Week 36<br>• Proportion of subjects with histologic remission at PCP Week 36<br>• Proportion of subjects with endoscopic improvement at PCP Week 36 |

| | | |
|---|---|---|
| | | • Proportion of subjects with mucosal healing at PCP Week 36 |

Abbreviations: BMI, body mass index; CRP, c-reactive protein; ITT, intent-to-treat; PCP, placebo-controlled period; RHI, Robarts Histopathology Index; UC, ulcerative colitis.

Subgroup analyses will only be performed for applicable endpoints (eg, subgroup analyses for histologic remission and mucosal healing will not be performed for the subgroup of baseline RHI lamina propria neutrophils and neutrophils in the epithelium subscores > 0 [yes, no]). Subgroup analyses for secondary and exploratory endpoints not listed in Table 5 may be performed.

As exploratory analyses, the impact of baseline Mayo score on treatment effects for primary and secondary efficacy endpoints may be investigated via logistic regression models that include covariates for treatment group, baseline Mayo score, and the interaction between treatment group and baseline Mayo score. Odds ratio estimates and 95% CIs for the treatment effect of each GB004 treatment group according to baseline Mayo score may be presented. Predicted probabilities by treatment group according to baseline Mayo score may also be presented. These exploratory analyses may also be performed to assess the impact of baseline Modified Mayo score.

## 8.5. Efficacy Assessments

### 8.5.1. Mayo Score

The Mayo score (adapted from Schroeder, 1987) Modified Mayo score, Partial Mayo score, Mayo subscores, and 6-point Mayo score are described in Protocol Section 8.1.2. Calculation of the Mayo score, Modified Mayo score, Partial Mayo score, and 6-point Mayo score is summarized in Table 6 below:

**Table 6: Calculation of Mayo Scores**

| Score | Calculation | Range | Applicable Analysis Visits |
|---|---|---|---|
| Mayo score | Sum of SFS, RBS, PGA, and endoscopic subscore | 0 - 12 | Baseline, PCP Week 12, PCP Week 36 |
| Modified Mayo score | Sum of SFS, RBS, and endoscopic subscore | 0 - 9 | Baseline, PCP Week 12, PCP Week 36 |
| Partial Mayo score | Sum of SFS, RBS, and PGA | 0 - 9 | Baseline, PCP Week 2, PCP Week 4, PCP Week 8, PCP Week 12, PCP Week 18, PCP Week 24, PCP Week 30, PCP Week 36, PCP Week 40 (Follow-up) |
| 6-Point Mayo score | Sum of SFS and RBS | 0 - 6 | Baseline, PCP Week 2, PCP Week 4, PCP Week 8, PCP Week 12, PCP Week 18, PCP Week 24, PCP Week 30, PCP Week 36, PCP Week 40 (Follow-up) |

Abbreviations: PCP, placebo-controlled period; PGA, Physician's Global Assessment; RBS, rectal bleeding subscore; SFS, stool frequency subscore.

For calculation of Mayo score, Modified Mayo score, and applicable efficacy endpoints based on these scores at the analysis visits of PCP Week 12 and PCP Week 36, the first PGA on or after the date of the applicable endoscopic subscore in the analysis visit window will be used as the PGA subscore, and the corresponding PGA date will be used to determine the days of daily subject eDiary scores to be utilized for calculation of the SFS and RBS. If there are multiple endoscopic subscores in the analysis visit window, calculation will be based on the endoscopic subscore closest to the target visit day. If multiple endoscopic subscores are equally close to the target visit day, then the latest assessment will be utilized for calculation.

For applicable efficacy endpoints, Mayo subscores comprising the Mayo score calculated at the analysis visits of PCP Week 12 and PCP Week 36 per the above will always be selected for inclusion in the analysis at the PCP Week 12 and PCP Week 36 analysis visits and will be ineligible for windowing to a different analysis visit window. Furthermore, Mayo subscores subsequent to those used for calculation at the analysis visits of PCP Week 12 and PCP Week 36 will be ineligible for windowing to analysis visits prior to PCP Week 12 and PCP Week 36, respectively.

For calculation of Partial Mayo score, 6-point Mayo score, PGA, SFS, and RBS Mayo subscores, and other applicable efficacy endpoints based on these scores at post-baseline analysis visits other than PCP Week 12 and PCP Week 36, the PGA closest to the target day will be used as the PGA subscore, and the corresponding PGA date will be used to determine the days of daily subject eDiary scores to be utilized for calculation of the SFS and RBS. If multiple PGA subscores are equally close to the target visit day, then the latest assessment will be utilized for calculation.

At each applicable analysis visit, SFS and RBS will each be calculated based on the average of the most recent 3 days with non-missing data within 14 days prior to the PGA date. If the PGA at a given visit is missing, then the date of another assessment at the given visit (eg, vital signs) may instead be used to determine the 14 day calculation timeframe. If the date for a given visit is completely missing, then the 14 day calculation timeframe may be determined relative to the target visit day. For the analysis visits of Baseline, PCP Week 12, and PCP Week 36, the day of and the day after flexible sigmoidoscopy/colonoscopy and the day(s) of bowel preparation, if performed, will be excluded from eligible days for calculation. Days selected for SFS and RBS calculation may be different based on eDiary data availability. SFS and RBS will be considered missing for subjects with fewer than 3 available days of eDiary data in the relevant calculation timeframe at a given visit.

Mayo scores and subscores will be presented in a by-subject data listing.

#### 8.5.1.1. Mayo Subscores

##### 8.5.1.1.1. Physician's Global Assessment

The PGA acknowledges the three other Mayo subscores, the subject's recollection of abdominal discomfort and general sense of wellbeing, and other observations, such as physical findings and the subject's performance status. The PGA is assessed on a 0-3 scale: 0 = Normal, 1 = Mild Disease, 2 = Moderate Disease, and 3 = Severe Disease, at the same applicable analysis visits as Partial Mayo score in Table 6 above.

#### 8.5.1.1.2. Stool Frequency and Rectal Bleeding Score

Daily eDiary scores will be collected from each subject to capture stool frequency and rectal bleeding. The SFS and RBS will be calculated at the same applicable analysis visits as Partial Mayo score in Table 6 above, using daily stool frequency scores and daily rectal bleeding scores.

The SFS is calculated as follows:

SFS = (average of the 3 daily stool frequency scores, rounded to the nearest integer) minus the reference normal stool frequency.

A subject's reference normal stool frequency is the number of stools in a 24-hour period when the subject was in remission from UC, or, if the subject does not report having previously achieved remission from UC, the number of stools the subject reports per day before initial onset of signs and symptoms of UC.

The SFS is assessed on a 0-3 scale: 0 = Normal number of stools for this subject; 1 = 1 to 2 stools more than normal; 2 = 3 to 4 stools more than normal; and 3 = 5 or more stools than normal.

The RBS is calculated as follows:

RBS = average of the 3 daily rectal bleeding scores, rounded to the nearest integer.

The RBS is assessed on a 0-3 scale: 0 = No blood seen; 1 = Streaks of blood with stool less than half the time; 2 = Obvious blood with stool most of the time; and 3 = Blood alone passes.

#### 8.5.1.1.3. Endoscopic Subscore

Endoscopic subscore of the sigmoid colon will be assessed by both a central reader (who is blinded to treatment assignment, visit, and subject symptoms) and a local endoscopist (who is blinded to treatment assignment). The same local endoscopist is to assess the endoscopic subscore for a particular subject throughout the study, if possible. Differences between the local endoscopist and the central reader will be adjudicated by a second blinded central reader as described in the study Imaging Charter.

Centrally read endoscopic subscores (from adjudication, if applicable) will be used for both subject eligibility determination and all applicable analyses.

The endoscopic subscore is assessed on a 0-3 scale: 0 = Normal or inactive disease; 1 = Mild disease (erythema, decreased vascular pattern); 2 = Moderate disease (marked erythema, absent vascular pattern, friability, erosions); and 3 = Severe disease (spontaneous bleeding, ulceration), at the same applicable analysis visits as the Mayo score in Table 6 above. An endoscopic subscore of 1 does not include friability; an endoscopy with friability will be assessed an endoscopic subscore of at least 2.

### 8.5.2. Histologic Assessments

#### 8.5.2.1. Robarts Histopathology Index

The RHI (Mosli, 2017) is described in Protocol Section 8.1.3.

RHI is evaluated by a blinded central reader for the sigmoid colon at Baseline, PCP Week 12, and PCP Week 36. The RHI score consists of four subscores: chronic inflammatory infiltrate,

lamina propria neutrophils, neutrophils in epithelium, and erosion or ulceration. The RHI score ranges from 0-33 and is derived from the Geboes score (Geboes, 2000), which is described in Section 8.5.2.2, as shown in Table 7 below.

The RHI score is calculated as follows:

RHI score = 1 x chronic inflammatory infiltrate level (4 levels)
+ 2 x lamina propria neutrophils (4 levels)
+ 3 x neutrophils in epithelium (4 levels)
+ 5 x erosion or ulceration (4 levels after combining Geboes 5.1 and 5.2)

**Table 7: RHI Scoring in Relation to Geboes Score**

| **RHI Subscore Subscore Categories** | **Numeric Value for RHI Calculation** | **Geboes Score/Subgrade** |
|---|---|---|
| Chronic inflammatory infiltrate | | 1 |
| No Increase | 0 | 1.0 |
| Mild but unequivocal increase | 1 | 1.1 |
| Moderate increase | 2 | 1.2 |
| Marked increase | 3 | 1.3 |
| Lamina propria neutrophils | | 2B |
| None | 0 | 2B.0 |
| Mild but unequivocal increase | 1 | 2B.1 |
| Moderate increase | 2 | 2B.2 |
| Marked increase | 3 | 2B.3 |
| Neutrophils in epithelium | | 3 |
| None | 0 | 3.0 |
| < 5% crypts involved | 1 | 3.1 |
| < 50% crypts involved | 2 | 3.2 |
| > 50% crypts involved | 3 | 3.3 |
| Erosion or ulceration | | 5 |
| No erosion, ulceration or granulation tissue | 0 | 5.0 |
| Recovering epithelium + adjacent inflammation | 1 | 5.1 |
| Probable erosion – focally stripped | 1 | 5.2 |
| Unequivocal erosion | 2 | 5.3 |
| Ulcer or granulation tissue | 3 | 5.4 |

Abbreviation: RHI, Robarts Histopathology Index.

RHI scores and subscores will be presented in a by-subject data listing.

#### 8.5.2.2. Geboes Score

The Geboes score is divided into 6 grades: architectural changes [grade 0], chronic inflammatory infiltrate [grade 1], lamina propria eosinophils and neutrophils [grade 2A and 2B], neutrophils in epithelium [grade 3], crypt destruction [grade 4] and erosions or ulcerations [grade 5]. The Geboes score is calculated as the highest grade with a corresponding subgrade score > 0, excluding Geboes Grade 2A, and ranges from 0 to 5.

Grading for the Geboes score is summarized in Table 8 below.

**Table 8: Geboes Score**

| Geboes Score Grade Description | Subgrade Scoring |
|---|---|
| Grade 0 – Architectural changes | |
| No abnormality | 0 |
| Mild abnormality | 1 |
| Mild or moderate diffuse or multifocal abnormalities | 2 |
| Severe diffuse or multifocal abnormalities | 3 |
| Grade 1 – Chronic inflammatory infiltrate | |
| No Increase | 0 |
| Mild but unequivocal increase | 1 |
| Moderate increase | 2 |
| Marked increase | 3 |
| Grade 2A – Lamina propria eosinophils | |
| No increase | 0 |
| Mild but unequivocal increase | 1 |
| Moderate increase | 2 |
| Marked increase | 3 |
| Grade 2B – Lamina propria neutrophils | |
| None | 0 |
| Mild but unequivocal increase | 1 |
| Moderate increase | 2 |
| Marked increase | 3 |
| Grade 3 – Neutrophils in epithelium | |
| None | 0 |
| < 5% crypts involved | 1 |
| < 50% crypts involved | 2 |
| > 50% crypts involved | 3 |

| Geboes Score Grade Description | Subgrade Scoring |
|---|---|
| Grade 4 – Crypt destruction | |
| None | 0 |
| Probable – local excess of neutrophils in part of crypt | 1 |
| Probable – marked attenuation | 2 |
| Unequivocal crypt destruction | 3 |
| Grade 5 – Erosions or ulcerations | |
| No erosion, ulceration or granulation tissue | 0 |
| Recovering epithelium + adjacent inflammation | 1 |
| Probable erosion – focally stripped | 2 |
| Unequivocal erosion | 3 |
| Ulcer or granulation tissue | 4 |

Geboes scores will be presented in a by-subject data listing.

### 8.5.3. UC-100 Index

The Ulcerative Colitis Index-100 (UC-100) score (Jairath, 2019). is calculated as follows:

UC-100 Index = 1 + (16 x SFS) + (6 x Endoscopic subscore) + RHI score.

The UC-100 Index ranges from 1 to 100, with higher scores indicating greater disease severity.

UC-100 Index will be presented in a by-subject data listing.

# 9. SAFETY EVALUATION

## 9.1. Adverse Events

Verbatim AE terms will be coded to an SOC and PT using MedDRA Version 24.0 or later.

All analyses of AEs will be based on the principle of treatment emergence. An AE is considered to be treatment-emergent as follows:

- PCP Week 12 Analysis:
  - For subjects who consent only to protocol version 1.0 and for subjects who consent to protocol version 2.0 and withdraw from PCP at or prior to PCP Week 12: AEs that start on or after the first dose of PCP study treatment.
  - For subjects who consent to protocol version 2.0 and do not withdraw from PCP at or prior to PCP Week 12: AEs that start on or after the first dose of PCP study treatment through the PCP Week 12 nominal visit (if date of PCP Week 12 nominal visit is non-missing) inclusive or through Day 85 (if date of PCP Week 12 nominal visit is missing) inclusive.

- PCP Final Analysis:
  - For subjects not dosed in the OLE: AEs that start on or after the first dose of PCP study treatment;
  - For subjects dosed in the OLE: AEs that start on or after the first dose of PCP study treatment and prior to the first dose of OLE study treatment.

In general, if the treatment emergence of an AE is not clear due to a missing or incomplete AE start date/time and/or missing start time for first dose of PCP study treatment, the AE will be considered to be treatment-emergent for the given analysis, unless the non-missing portions of the AE start date/time and the end date/time and the date of first dose of PCP study treatment indicate otherwise.

For subjects who consent to version 2.0 of the protocol, in cases where an AE has a non-missing start date equal to the date of PCP Week 12, the AE will be considered treatment emergent for the PCP Week 12 Analysis if it is reported on the appropriate AE eCRF intended to capture AEs through PCP Week 12. If the date of PCP Week 12 is missing, then the target day for PCP Week 12 (Day 85) will be utilized for the purpose of determining treatment emergence.

In cases where an AE has a non-missing start date equal to the date of first dose of OLE study treatment, but the AE start time and/or time of first dose of OLE study treatment is missing, the AE will be considered treatment-emergent to the PCP if it is reported on the appropriate AE eCRF intended to capture AEs in the PCP.

In general, whenever a tabular summary of AEs is mentioned in this SAP, it is intended that the tabular summary is in reference to treatment emergent AEs, even though “treatment emergent” may not be explicitly mentioned. By-subject data listings of AEs will include all events regardless of treatment emergence.

Counting will be at the subject level for each level of summarization (eg, any AE, SOC, and PT), with subjects experiencing more than one AE counted only once. For summary of AEs by severity, subjects will be counted once at the highest severity reported at each level of summarization. AEs that are missing severity will be presented in summary tables as “Missing”. AEs that are missing relationship to study treatment will be presented in summary tables as “Related”.

The following AE summaries will be presented:

- Overall Summary of AEs, showing the number and percentage of subjects with an AE; a moderate or severe AE; a severe AE; a serious adverse event (SAE) including UC SAEs (defined as the PT of ‘Colitis ulcerative’); an SAE excluding UC SAEs; a treatment-related AE; a treatment-related SAE; an AE leading to dose reduction, dose interruption, or discontinuation of PCP study treatment; an AE leading to dose reduction; an AE leading to dose interruption; an AE leading to discontinuation of PCP study treatment; an AE leading to withdrawal from PCP; and an AE resulting in death.
- Incidence of AEs by SOC and PT
- Incidence of AEs by PT

- Incidence of AEs ≥ 5% in Any Treatment Group by PT
- Incidence of AEs ≥ 2% Higher in Either GB004 Treatment Group Relative to Placebo by PT
- Incidence of SAEs by SOC and PT including UC SAEs
- Incidence of SAEs by SOC and PT excluding UC SAEs
- Incidence of AEs by SOC, PT, and maximum severity
- Incidence of Treatment-Related AEs by SOC and PT
- Incidence of AEs Leading to Dose Reduction of PCP Study Treatment by SOC and PT
- Incidence of AEs leading to Dose Reduction of PCP Study Treatment by PT
- Incidence of AEs Leading to Dose Interruption of PCP Study Treatment by SOC and PT
- Incidence of AEs Leading to Dose Interruption of PCP Study Treatment by PT
- Incidence of AEs Leading to Discontinuation of PCP Study Treatment by SOC and PT
- Incidence of AEs Leading to Discontinuation of PCP Study Treatment by PT
- Incidence of AEs Leading to Withdrawal from PCP by SOC and PT
- Incidence of AEs Leading to Withdrawal from PCP by PT
- Incidence of AEs Resulting in Death by SOC and PT

In-depth summaries of nausea, dizziness, and/or other specific AEs may be presented, including information such as AE severity, treatment-related AEs, AEs requiring treatment, AEs leading to dose reduction, dose interruption, or discontinuation of PCP study treatment, AEs leading to withdrawal from PCP, onset timing, incidence over time, AE duration, prevalence over time, and duration of AE after last dose of PCP study treatment.

All AEs, AEs leading to dose reduction of PCP study treatment, AEs leading to dose interruption of PCP study treatment, AEs leading to discontinuation of PCP study treatment and/or withdrawal from PCP, AEs resulting in death, and SAEs will be presented in by-subject data listings, regardless of treatment emergence, with those AEs that are treatment-emergent identified. A by-subject listing of SAEs occurring in screen failures may be presented, if any occur.

## 9.2. Clinical Laboratory Evaluation

Actual values, change from baseline values, and percent change from baseline values for quantitative laboratory parameters (hematology, clinical chemistry, plasma markers [erythropoietin (EPO) and vascular endothelial growth factor (VEGF)], urinalysis, and serum markers [iron, transferrin, total iron binding capacity, serum ferritin, hepcidin, and zinc]) will be summarized by visit.

The following quantitative clinical chemistry laboratory parameters will be derived based on reported laboratory parameters:

- Indirect bilirubin, derived by subtracting direct bilirubin from total bilirubin
- Estimated glomerular filtration rate (eGFR), derived using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) 2021 equation (Inker, 2021):

  eGFR (ml/min/1.73m$^2$) =
  142 x minimum (SCr/k, 1)$^{\alpha}$ x maximum (SCr/k, 1)$^{-1.200}$ x 0.9938$^{Age}$ x
  1.012 [if female],

  where SCr is serum creatinine (in mg/dL), k is 0.7 for females and 0.9 for males, α is -0.241 for females and -0.302 for males, minimum (SCr/k, 1) indicates the lesser of SCr/k and 1, and maximum indicates the greater of SCr/k and 1, and Age is baseline age in years.

Shift tables will be presented, summarizing shifts from baseline to high and low or shifts from baseline to abnormal/positive at any time post-baseline based on laboratory normal ranges.

Summaries of maximum post-baseline values for liver chemistry parameters (alanine aminotransferase [ALT], aspartate aminotransferase [AST], ALT or AST, total bilirubin, concurrent ALT or AST and total bilirubin, total bilirubin, direct bilirubin, indirect bilirubin, alkaline phosphatase, gamma-glutamyl transferase [GGT]) according to categories based on multiples of upper limit of normal (ULN) will be presented. Liver chemistry parameters according to categories based on multiples of ULN may also be presented by visit.

The number and percentage of subjects with an increase from baseline in hemoglobin to a high (> ULN) value at any time post-baseline will be presented for the following categories: > 0 g/L, > 20 g/L, and > 40 g/L.

In addition, the following figures/plots will be provided:

- Mean actual, change from baseline, and percent change from baseline values, including SE bars, by visit figures for select laboratory parameters (eg, hemoglobin, creatinine, eGFR, EPO, and VEGF)
- Evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) (Merz, 2014) plots for maximum ALT, maximum AST, and maximum ALT or AST versus total bilirubin

Creatinine and eGFR may be further assessed in the subset of subjects with concomitant use of sensitive clinical substrates of organic cation transporter 2 (such as metformin or dalfampridine).

Laboratory parameters will also be presented in by-subject data listings.

## 9.3. Vital Signs

Actual values and change from baseline values for vital sign parameters (systolic and diastolic blood pressure, heart rate, respiratory rate, temperature, and weight) will be summarized by visit.

A summary table of the incidence of abnormalities in vital sign parameters will be presented according to the abnormality criteria listed below in Table 9.

**Table 9: Vital Sign Parameter Abnormality Criteria**

| Vital Sign Parameter | Abnormality Criteria |
|---|---|
| Systolic Blood Pressure | High or Increased: > 180 mmHg post-baseline if ≤ 180 mmHg at baseline, or an increase from baseline of > 40 mmHg.<br>Low or Decreased: < 90 mmHg post-baseline if ≥ 90 mmHg at baseline, or a decrease from baseline of > 30 mmHg. |
| Diastolic Blood Pressure | High or Increased: > 105 mmHg post-baseline if ≤ 105 mmHg at baseline, or an increase from baseline of > 30 mmHg.<br>Low or Decreased: < 50 mmHg post-baseline if ≥ 50 mmHg at baseline, or a decrease from baseline of > 20 mmHg. |
| Pulse Rate | High or Increased: > 120 beats per minute (bpm) post-baseline if ≤ 120 bpm at baseline, or an increase from baseline of > 20 bpm.<br>Low or Decreased: < 50 bpm post-baseline if ≥ 50 bpm at baseline, or a decrease from baseline of > 20 bpm. |
| Temperature | > 38 degrees C and an increase from baseline of ≥1 degree C. |

Abbreviations: bpm, beats per minute; C, Celsius; mmHg, millimeters of mercury.

Vital sign parameters will also be presented in a by-subject data listing, with abnormalities flagged.

## 9.4. 12-Lead ECGs

Actual values and change from baseline values will be summarized by visit for the following quantitative ECG parameters: mean heart rate; RR interval; QT interval (uncorrected); Fridericia's correction formula for QT interval (QTcF); PR interval; and QRS duration.

Outlier analyses for QTcF intervals in the PCP will be performed. This will consist of a summary of the number and percentage of subjects with a post-baseline QTcF interval greater than 450 msec, 480 msec, and 500 msec and the number and percentage of subjects with an increase from baseline in QTcF interval of greater than 30 msec and 60 msec.

Each ECG will be assessed by a blinded central reader with an overall interpretation of normal, abnormal, or unable to evaluate. Shift tables for overall interpretation will be presented, summarizing shifts from baseline to abnormal at any time post-baseline.

ECG parameters will also be presented in a by-subject data listing with values meeting the above QTcF outlier criteria and values representing a shift from baseline to abnormal in overall interpretation flagged.

# 10. CHANGES TO THE ANALYSES PLANNED IN THE PROTOCOL

There are no changes to the analyses planned in the protocol.

# 11. REFERENCES

Chen J, Hunter S, Kisfalvi K, Lirio RA. A hybrid approach of handling missing data under different missing data mechanisms: VISIBLE 1 and VARSITY trials for ulcerative colitis. *Contemp Clin Trials.* 2021;100:106226.

EMA, 2020. *Points to consider on implications of Coronavirus disease (COVID-19) on methodological aspects of ongoing clinical trials*. Committee for Medicinal Products for Human Use (CHMP). EMA/158330/2020 Rev. 1. 26 June 2020. Amsterdam, The Netherlands. https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-implications-coronavirus-disease-covid-19-methodological-aspects-ongoing-clinical_en-0.pdf

EMA, 2021. *Guidance on the management of clinical trials during the COVID-19 (Coronavirus) pandemic*. Version 4 (04/02/2021) February 2021. Amsterdam, The Netherlands. https://ec.europa.eu/health/system/files/2022-02/guidanceclinicaltrials_covid19_en_1.pdf

FDA, 2020a. *FDA Guidance on conduct of clinical trials of medical products during COVID-19 pandemic*. Guidance for Industry, Investigators, and Institutional Review Boards. March 2020, Updated 30 August 2021. https://www.fda.gov/media/136238/download

FDA, 2020b. *Statistical considerations for clinical trials during the COVID-19 public health emergency*. Guidance for Industry. June 2020. https://www.fda.gov/media/139145/download

Geboes K, Riddell R, Ost A, Jensfelt B, Persson T, Lofberg R. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. *Gut.* 2000;47(3):404-409.

ICH E3, 1995. Harmonised Tripartite Guideline: *Structure and content of clinical study reports E3*. Step 4, 1995. (n.d.). https://database.ich.org/sites/default/files/E3_Guideline.pdf

ICH E9, 1998. Harmonised Tripartite Guideline: *Statistical principles for clinical trials E9*. Step 4, 1998. (n.d.). https://database.ich.org/sites/default/files/E9_Guideline.pdf

ICH E9(R1), 2019. Harmonised Guideline: *Addendum on estimands and sensitivity analysis in clinical trials To the Guideline on statistical principles for clinical trials E9(R1)*. Step 4, 2019. (n.d.). https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

Inker LA, Eneanya ND, Coresh J, et al. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. *N Engl J Med.* 2021;385(19):1737-1749.

Jairath V, Jeyarajah J, Zou G, et al. A composite disease activity index for early drug development in ulcerative colitis: development and validation of the UC-100 score. *The Lancet Gastroenterology & Hepatology.* 2019;4(1):63-70.

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst.* 1959;22(4):719-748.

Mantel N, Fleiss JL. Minimum expected cell size requirements for the Mantel-Haenszel one-degree-of-freedom chi-square test and a related rapid procedure. *Am J Epidemiol.* 1980;112(1):129-134.

Merz M, Lee KR, Kullak-Ublick GA, Brueckner A, Watkins PB. Methodology to assess clinical liver safety data. *Drug Saf.* 2014;37 Suppl 1:S33-45.

Mosli MH, Feagan BG, Zou G, et al. Development and validation of a histological index for UC. *Gut.* 2017;66(1):50-58.

Newcombe RG. Interval estimation for the difference between independent proportions: comparison of eleven methods. *Statist Med.* 1998.

Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med.* 1987;317(26):1625-1629.

Tarone R. On heterogeneity tests based on efficient scores. *Biometrika.* 1985;72(1):91-95.

Wilson EB. Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association.* 1927;22(158):209-212.