STATISTICAL ANALYSIS PLAN

A PHASE 3, MULTI-CENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED TRIAL OF ORAL RPC1063 AS INDUCTION AND MAINTENANCE THERAPY FOR MODERATE TO SEVERE ULCERATIVE COLITIS

INVESTIGATIONAL PRODUCT:	**RPC1063**
PROTOCOL NUMBER:	**RPC01-3101**
PROTOCOL VERSION #1.0 FINAL:	**3/30/2015**
PROTOCOL VERSION #2.0 FINAL:	**6/7/2016**
PROTOCOL VERSION #3.0 FINAL:	**6/7/2017**
PROTOCOL VERSION #4.0 FINAL:	**12/7/2017**
PROTOCOL VERSION #5.0 FINAL:	**5/29/2018**
PROTOCOL VERSION #6.0 FINAL:	**11/26/2018**
PROTOCOL VERSION #7.0 FINAL:	**5/20/2019**

HISTORY OF AMENDMENT(S)

Amendment number	Date of amendment	Significant amendments and rationales	Impact to statistical analysis and descriptions
V4.0	6/18/2019	• Changes to the SAP are made based on FDA's response letter (ADVICE/INFORMATION REQUEST) dated on 15 June, 2019. They include: 1. Primary analysis algorithm: The 7-day scoring algorithm as described in Appendix 3.2 is the primary analysis algorithm for primary and key secondary endpoints (clinical remission and response). 2. Analysis algorithm for sensitivity analysis for clinical remission and response: • The 14-day scoring algorithm described in Appendix 3.1 will be used for the sensitivity analysis of clinical remission and response. This is called "14-day algorithm A". • An additional "14-day scoring algorithm B" (Appendix 3.1) for sensitivity analysis is based on the following: -For calculation of stool frequency (SF) and rectal bleeding (RB) subscores, 3 days are required. -If data from 3 consecutive days are not available, then data from 2 consecutive days and a non-consecutive day closest to the visit date (without including the day(s) before, day of, or day(s) after endoscopy) would be acceptable. -Both SF and RB data must be available on the same day. • Simplify treatment compliance algorithm described in Sec 9.3. • Include moderate UC at screening in subgroup analysis.	• The 7-day scoring algorithm will be used in the primary analyses of primary and key secondary endpoints (clinical remission and response), while the 14-day scoring algorithm (A and B) will serve as sensitivity analyses of these two endpoints.
V3.0	3/20/2019	• Adopted Celgene's current SAP template. • Aligned the SAP content to match the current protocol amendment 6 and keep consistency in data handling	• The analyses of primary and key secondary endpoints (clinical remission and response) [redacted]

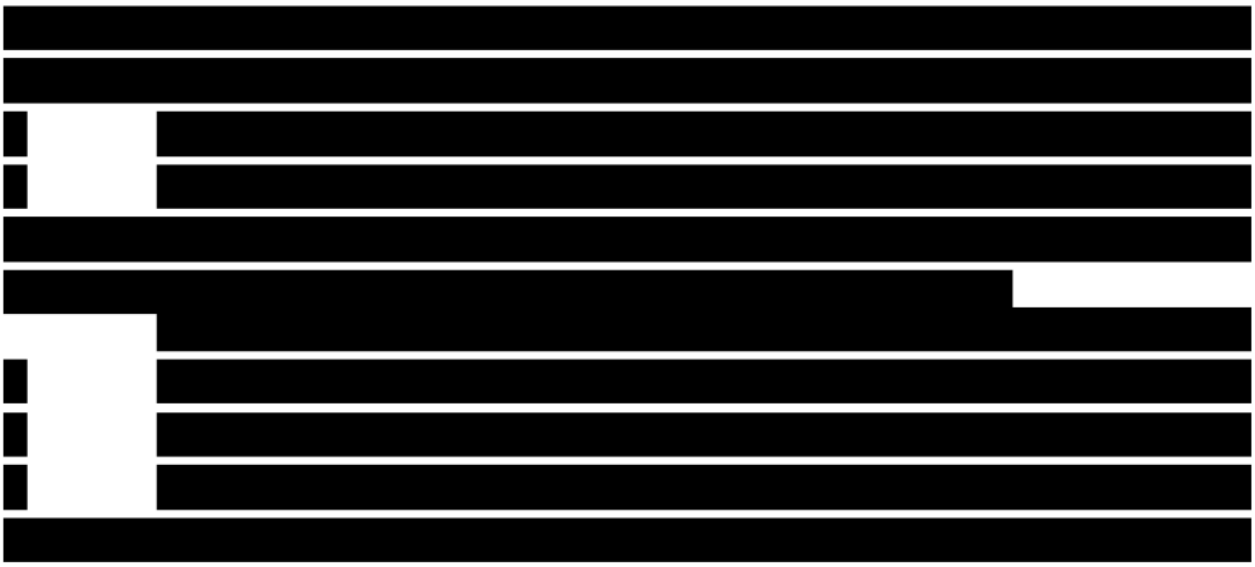
		(including missing data handling) with the induction period SAP. • Section 2. Clarification on timing of Cohort 1 analysis. • Section 3. Clarification on dosing compliance rates. • Section 4.1. Adjustment of Cohort 1 and Cohort 2 sample sizes according to the current protocol. [redacted] • Section 5.4. Updated the list of reasons to exclude patients from the PP population to align with the protocol. • Section 6.1.2, Section 6.2.3, Section 6.3.2. Sensitivity and additional analyses were were included to support the primary analysis of the primary and key secondary efficacy endpoints. • Section 6.2. Consolidation of analysis description of all 3 secondary efficacy endpoints. • Section 6.4. Removed analysis of Mayo score using local endoscopy and certain subgroup analysis due to insufficient data. Include additional subgroup analyses based on FDA's recommendation (FDA Type C Meeting Request - Written Responses letter received on 4 February 2019). • Appendix 1.3. Non-responder imputation for anyone with any missing Mayo subscore will be used as the primary analysis method of the primary endpoint. Removal of mean observed value (MOV) imputation to be in line with other Celgene program. [redacted] • Appendix 1.3. Analysis visit window was revised to to facilitate analyses of efficacy data. [redacted] • Protocol-specified summary of total and average daily doses was removed due to limitation of dosing log CRF design. • Incorporated comments from FDA Type C Meeting Request - Written	[redacted] • The removal of imputation in the event of a single missing Mayo subscore allows a more objective assessment of efficacy at the timepoint of interest. • The sensitivity analysis based on PP population will be aligned with in PP criteria in the current SAP. • The sensitivity analyses using tipping point analysis and multiple imputation for missing data handling for the clinical remission and response endpoints will enable a robustness assessment of the primary analysis of the two endpoints. • Additional subgroup analyses will enable us assess the consistency of efficacy results in various subgroups. • New analysis window will allow more data to contribute to the analysis so it will help address some missing data issue. [redacted] • Removal of MOV imputation should have low impact on the interpretation of the efficacy analysis results, since it's used in one of the sensitivity analyses.

		Responses letter received on 4 February 2019.	
V2.0	9/25/2015	• Section 1.1. Sponsor and Oversight: Changed Statistical analysis CRO to PPD, LLC • Section 2. Introduction: Modified the language regarding the final analyses and data unblinding/lock for IP [redacted] • Section 5.6.5. Handling of Missing and Partial Data: Changed the language to eliminate the imputation of mayo subscore. [redacted]	
V1.0 (Original)	5/29/2015	Initial version	

SIGNATURE PAGE

STATISTICAL ANALYSIS PLAN (SAP) AND SAP AMENDMENT APPROVAL SIGNATURE PAGE	
SAP TITLE	Statistical Analysis Plan (Induction Period, Cohort 1 and Cohort 2)
SAP VERSION, DATE	Version 4.0, 6/18/2019
[redacted]	
PROTOCOL TITLE	A PHASE 3, MULTI-CENTER, RANDOMIZED, DOUBLE-BLIND, PLACEBO CONTROLLED TRIAL OF ORAL RPC1063 AS INDUCTION AND MAINTENANCE THERAPY FOR MODERATE TO SEVERE ULCERATIVE COLITIS
INVESTIGATIONAL PRODUCT	RPC1063
PROTOCOL NUMBER	RPC01-3101
PROTOCOL VERSION, DATE	Version 7.0, 05/20/2019
SIGNATURE STATEMENT	By my signature, I indicate I have reviewed this SAP and find its contents to be acceptable.
[redacted]	
[redacted]	

TABLE OF CONTENTS



LIST OF TABLES

LIST OF FIGURES

1. LIST OF ABBREVIATIONS

Table 1: Abbreviations and Specialist Terms

Abbreviation	Term
5-ASA	5-aminosalicylic acid
6-MP	6-mercaptopurine
aCRF	annotated case report form
ADaM	analysis data model
AE	adverse event
AESI	adverse event of special interest
ALC	absolute lymphocyte count
ALT	alanine aminotransferase
ANC	absolute neutrophil count
ANCOVA	analysis of covariance
anti-HBc total	anti-hepatitis B core total antibodies
anti-TNF	anti-tumor necrosis factor
AST	aspartate aminotransferase
ATC	anatomic therapeutic classification
AV	atrioventricular
BMI	body mass index
BUN	blood urea nitrogen
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRP	C-reactive protein
CRO	contract research organization
DBP	diastolic blood pressure
DLCO	diffusion capacity of the lung for carbon monoxide
DM	data management
DNA	deoxyribonucleic acid

DSMB	data safety monitoring board
ECG	electrocardiogram
eCRF	electronic case report form
EQ-5D	EuroQol quality of life scale
FDA	Food and Drug Administration
FEV_1	forced expiratory volume in 1 second
FVC	forced vital capacity
GGT	gamma glutamyltransferase
hCG	human chorionic gonadotropin
HAV	hepatitis A virus
HBsAG	surface antigen of the hepatitis B virus
HBV	hepatitis B virus
HCV	hepatitis C virus
HDL	high-density lipoprotein
HIV	human immunodeficiency virus
HR	heart rate
IG	implementation guide
IgG	immunoglobulin G
IgM	immunoglobulin M
IP	induction period
ITT	intent-to-treat
IXRS	interactive voice/web-based activated response system
LDL	low-density lipoprotein
LFT	liver function test
MCH	mean corpuscular hemoglobin
MCHC	mean corpuscular hemoglobin concentration
MCV	mean corpuscular volume
MedDRA	medical dictionary for regulatory activities

MI	multiple imputation
MP	maintenance period
ms	milliseconds
NRI	non-responder imputation
OCT	optical coherence tomography
OLE	open-label extension
PD	pharmacodynamic(s)
PFT	pulmonary function test
PGA	physician's global assessment or physician's global assessment subscore
PK	pharmacokinetic(s)
PP	per protocol
PT	preferred term
QOL	quality of life
RBC	red blood cell
RBS	rectal bleeding subscore
S1P	sphingosine-1-phosphate
S1P1R	sphingosine-1-phosphate 1 receptor
S1P5R	sphingosine-1-phosphate 5 receptor
SAE	serious adverse event
SAP	statistical analysis plan
SBP	systolic blood pressure
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SF-36	short form 36 health survey
SF-6D	health utility score derived from the SF-36
SFS	stool frequency subscore
SI	standard international

SOC	system organ class
TEAE	treatment-emergent adverse event
TFR	treatment failure rules
TNF	tumor necrosis factor
TPA	tipping point analysis
UC	ulcerative colitis
UK	United Kingdom
ULN	upper limit of normal
US	United States
VAS	visual analogue scale
VZV	varicella-zoster virus
WBC	white blood cell
WHO DDE	World Health Organization drug classification dictionary enhanced
WPAI:UC	work productivity and activity impairment questionnaire for UC

2. INTRODUCTION

Ulcerative colitis (UC) is a chronic gastrointestinal inflammatory disorder that involves the surface mucosa, the crypt epithelium, and the submucosa of the colon ([redacted] The etiology of UC is multifactorial, but likely includes a dysregulated mucosal immune response against commensal non-pathogenic bacteria of the colon, resulting in bowel inflammation [redacted] The overall goal of treatment for patients with active UC is to induce and maintain remission and to induce and maintain mucosal healing [redacted] Treatment of UC consists of anti-inflammatory and immunosuppressive therapies that are chosen to maximize efficacy while avoiding toxicity. There remains an unmet need for a UC treatment that is highly effective, well-tolerated, and orally active in both the adult and pediatric populations.

RPC1063 is a small molecule compound that selectively and potently activates the sphingosine-1-phosphate 1 receptor ($S1P_1$) and the S1P 5 receptor ($S1P_5$), although it is more selective towards $S1P_1$ over $S1P_5$.

The induction phase of a Phase 2 trial in adult patients with moderate to severe UC (RPC01-202) was completed in October 2014. At the conclusion of the induction phase, the proportion of patients achieving clinical response and clinical remission with RPC1063 1 mg was greater than placebo and the difference was both clinically meaningful and statistically significant. In addition, all secondary endpoints at the conclusion of the induction phase, including clinical response, change in the Mayo score, and mucosal improvement on endoscopy, were also positive and statistically significant for the RPC1063 1 mg dose. The maintenance phase of this trial ended March 2015 and an open-label extension period is ongoing.

The objective of the RPC1063 clinical development program in UC is to demonstrate that RPC1063 administered orally is safe and effective in inducing and maintaining remission in patients with moderate to severe UC. Positive Phase 2 results on remission rates during the Induction Period of protocol RPC01-202 suggest that RPC1063 has the potential to be a clinically meaningful addition to the therapeutic armamentarium for the treatment of moderate to severe UC.

The results from this trial are meant to serve as the primary efficacy and safety data to support an application for marketing approval from regulatory authorities.

For the purposes of statistical analyses, the induction period (IP) and the maintenance period (MP) will be treated as two independent trials thus each analysis will retain the full alpha. This statistical analysis plan (SAP) presents a detailed plan of the statistical methods to be used during the reporting and analysis of clinical data collected in the IP of this trial in adult patients only. The adult patient data from the IP in Cohort 1 will be cleaned, unblinded and analyzed after the last Cohort 1 adult patient has had their last IP visit (I 3 Visit). The analysis of Cohort 1

induction data which is used for planning for future development will be carried out by designated un-blinded Biostatistics and Programming teams at PPD (a CRO), the same teams already contracted to support the on-going DSMB meetings for this study. A limited number of Sponsor (i.e., Celgene) members who are not part of the study team will have access to the Cohort 1 induction analysis results. The review and conduct of the study will continue in a blinded manner by study team members that are blinded to all study data until all randomized subjects have completed 42 weeks of the Maintenance Phase (Week 52 of overall study duration). The process the Cohort 1 induction analysis will be documented in a charter in a separate document.

The adult patient data from Cohort 2 will be cleaned, unblinded and summarized after the last Cohort 2 adult patient has had their last Cohort 2 visit. Analysis of all clinical data collected for adult patients who enter the MP will be described in a separate SAP. Analysis of clinical data collected for pediatric [redacted]

This plan should be read in conjunction with the trial protocol, the annotated case report forms (aCRFs), and vendor data transfer specifications.

3. STUDY OBJECTIVES

Induction Therapy

The primary objective is to:

- Demonstrate the efficacy of RPC1063 versus placebo on induction of clinical remission in adults

The secondary objectives are to:

- Demonstrate the efficacy of RPC1063 versus placebo on induction of clinical response in adults
- Demonstrate the efficacy of RPC1063 versus placebo on achieving endoscopic improvement in adults
- [redacted]
- [redacted]

4. INVESTIGATIONAL PLAN

4.1. Overall Study Design and Plan

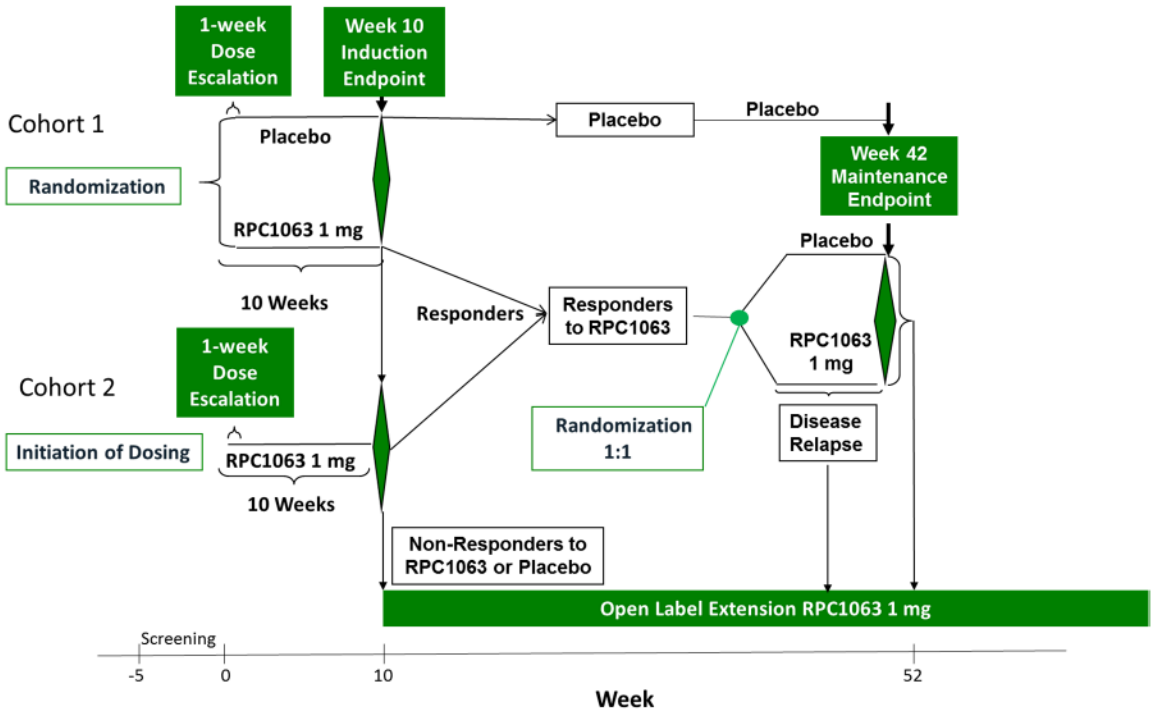
This is a multicenter, randomized, double-blind, placebo-controlled trial of RPC1063 as induction and maintenance therapy for moderate to severe UC. The trial will be conducted at approximately 325 sites in North America, Europe, Asia Pacific, South America, South Africa.

The trial is composed of 2 periods: Induction and Maintenance. Adult patients will enter into the trial through the Induction period (IP) in 2 separate cohorts (see below). Adult patients who have previously received anti-TNF therapy may commence enrollment into Cohort 2 once the randomization limit of approximately 30% anti-TNF therapy adult patients has been reached in Cohort 1. Those adult patients who are anti-TNF therapy naïve will continue to enroll into Cohort 1 and can enroll into Cohort 2 only after Cohort 1 has been closed to enrollment. The proportion of adult patients who have previously received anti-TNF therapy will be limited to approximately 50% in Cohort 2.

Adult patients from Cohort 1 or Cohort 2 with clinical response at the end of the IP will proceed through to the MP. Patients who participate in this trial may participate in an optional Open-Label Extension (OLE) trial as described below.

Approximately 900 adult patients will be entered into the trial in 2 separate cohorts through the IP, and may proceed through to the MP and/or the OLE trial.

Figure 1: Trial Schematic for Cohort 1 and Cohort 2



Induction Period

The IP of the trial is composed of 2 cohorts that will be treated for 10 weeks and evaluated for clinical response/remission:

- Cohort 1: Approximately 600 adult patients will be randomized in a 2:1 ratio to receive either RPC1063 1 mg (400 patients) or placebo (200 patients) once daily in a double-blind fashion, stratified by corticosteroid use at Screening (yes or no), and prior anti-TNF use (yes or no)
- Cohort 2: Approximately 300 adult patients* will receive open-label RPC1063 1 mg once daily

*The number of adult patients in Cohort 2 may be increased if necessary to ensure approximately 400 patients who have received RPC1063 will be randomized in the MP.

Patient eligibility for the IP will be determined during a 5-week Screening Period prior to entry.

The trial will include both patients that have received anti-TNF therapy and those who have not. The proportion of adult patients who have previously received anti-TNF therapy will be limited to approximately 30% in Cohort 1. The proportion of adult patients who have previously received anti TNF therapy will be limited to approximately 50% in Cohort 2.

All adult patients will initiate investigational drug in accordance with a 7-day dose escalation regimen starting with RPC1063/ozanimod HCl 0.25 mg (equivalent to ozanimod 0.23 mg) or matching placebo (matching placebo for Cohort 1 only) on Days 1 to 4 and RPC1063/ozanimod HCl 0.5 mg (equivalent to ozanimod 0.46 mg) once daily or matching placebo on Days 5 to 7. On Day 8, patients will receive the final dose level RPC1063/ozanimod HCl 1 mg (equivalent to ozanimod 0.92 mg) once daily or matching placebo for 9 weeks.

The total duration of the IP is 10 weeks.

Maintenance Period

It is anticipated that approximately 400 adult patients treated with RPC1063 during the IP and who complete the IP (both Cohort 1 and Cohort 2) will be in clinical response (by either definition in Section 4.2.1) at Week 10 and will be eligible to enter the randomized, double-blind, placebo-controlled MP. These adult patients in clinical response at Week 10 of the IP will be randomized to receive either RPC1063 1 mg or matching placebo in a 1:1 ratio. who received RPC1063 during Induction will be randomized to receive either RPC1063 1 mg or matching placebo in a 1:1 ratio. Adult patients in clinical response at Week 10 of the IP who were randomized to placebo (Cohort 1) will continue to receive placebo in the MP in a double-blind manner. Adult patients from Cohort 1 or Cohort 2 who do not show a clinical

response at Week 10 or who lose response at any time during the MP may enter the optional OLE trial, if appropriate.

Adult patients randomized in the MP will be stratified by clinical remission status (by either definition in Section 4.2.1) at Week 10 (yes or no) and corticosteroids use at Week 10 (yes or no). Adult patients will be evaluated for disease activity/efficacy at Week 42 of the MP (52 weeks total treatment).

Adult patients who complete the IP and are non-responders at Week 10, or those who complete the MP, or those who experience disease relapse during the MP, will have the option to enter a separate OLE trial (RPC01-3102).

4.2. Study Endpoints

4.2.1. Definitions

Complete Mayo score: the sum of the Rectal Bleeding subscore, Stool Frequency subscore, Physician Global Assessment subscore, and the Endoscopy subscore. Each subscore has a range of 0-3 points and the complete Mayo score has a range of 0-12 points

The Mayo endoscopic subscore will be determined both at the investigator site and centrally [redacted]

Clinical Remission

Three-component Mayo: Rectal Bleeding subscore = 0 and Stool Frequency subscore ≤ 1 (and a decrease of ≥ 1 point from the Baseline Stool Frequency subscore) and Endoscopy subscore ≤ 1

Four-component Mayo: Complete Mayo score of ≤ 2 points with no individual subscore of > 1 point

Clinical Response

Three-component Mayo: A reduction from Baseline in the 9-point Mayo score of ≥ 2 points and ≥ 35%, and a reduction from Baseline in the Rectal Bleeding subscore of ≥ 1 point or an absolute Rectal Bleeding subscore of ≤ 1 point

[redacted]

Durable Clinical Remission: Clinical remission at Week 10 and at 52 weeks in all patients who entered the MP

Maintenance of Remission: Clinical remission at 52 weeks in the subset of patients who are in remission at Week 10

Corticosteroid-free Remission: Clinical remission at 52 weeks while off corticosteroids for ≥ 12 weeks

Endoscopic Improvement: Endoscopy subscore of ≤ 1 point

Mucosal Healing: Endoscopy subscore of ≤ 1 point and a Geboes index score < 2.0

[redacted]

Treatment Failure:

Adult patients will be considered to have failed treatment if any of the following occur:

- Any protocol-prohibited change in medications including:
  - Post-Baseline initiation of, or increase in total daily dose level higher than the maximum dose taken between the Screening and Baseline visits in:
    - Corticosteroids or 5-ASA dose to treat UC
    - Prolonged course of systemic corticosteroids > 7 days for treatment of disease other than UC
  - Initiation of an immune suppressing therapy including 6-mercaptopurine (6-MP), azathioprine, anti-TNF agents, or vedolizumab
- A colectomy (partial or total) or an ostomy
- Discontinuation of investigational drug for lack of therapeutic effect before the Week 10 or 52 week efficacy evaluations

Treatment failure will be confirmed via a blinded clinical data review process as per Appendix 1.3.

4.2.2. Endpoints

Efficacy - Cohort 1

The primary efficacy endpoint is:

- The proportion of adult patients in clinical remission at Week 10

The key secondary efficacy endpoints are:

- The proportion of adult patients with a clinical response at Week 10
- The proportion of adult patients with endoscopic improvement at Week 10
- The proportion of adult patients with mucosal healing at Week 10

[redacted]

Cohort 2

Cohort 2 is open-label; therefore no formal analysis of efficacy endpoints will be conducted. All efficacy endpoints listed under Cohort 2 will be summarized and described without hypothesis testing and reported using descriptive statistics.

[redacted]

5. DEFINITION OF ANALYSIS POPULATIONS

Unless specified otherwise, patients refer to adult patients from Cohort 1 and/or Cohort 2 only.

5.1. Intent-to-Treat (ITT) Population

The ITT population will consist of:

- Cohort 1: All randomized patients from Cohort 1 of the IP of the trial who received at least 1 dose of investigational drug (RPC1063 or placebo)
- Cohort 2: All enrolled patients from Cohort 2 of the IP of the trial who received at least 1 dose of investigational drug (RPC1063)

The ITT population will be used as the primary population for all efficacy parameters. Patients who prematurely withdrew from the trial for any reason and for whom an assessment was not performed for any reason will still be included in the ITT analyses. Patients in Cohort 1 who received an incorrect investigational drug from that which was randomized will be summarized in the treatment group according to the intended randomization. Patients in Cohort 1 who were randomized with a misreported stratum will be analyzed according to their original [misreported] stratum.

Patients receiving RPC1063 in Cohort 2 will be reported in either a separate column from patients randomized to RPC1063 in Cohort 1 or in a separate table and will not be subject to statistical hypothesis testing.

5.2. Per-Protocol (PP) Population

The PP population will consist of all patients in the ITT population who adhered to the protocol with no major protocol deviation that potentially affects efficacy in the induction period. This population will be used in the sensitivity analysis of clinical remission and additional analysis of clinical response to evaluate the influence of major protocol violators and protocol deviators on the primary results. Patients will be excluded from this population if they meet any of the following reasons, which will be documented prior to database snapshot:

- Investigational drug noncompliance > 20% (i.e., overall IP drug compliance not in [80%, 120%])
- Receiving incorrect investigational drug leading to a treatment change for more than 1 week in the IP
- Prolonged concomitant use of systemic corticosteroids (except topical use) > 7 days before Week 10, for treatment of disease other than UC, at a total daily dose higher than the maximum dose taken between the Screening and Baseline

- Concomitant use of protocol-specified excluded medication that potentially affects efficacy for unrelated comorbid condition

[redacted]

- Violation of key entry criteria (i.e., inclusion criterion #4 [Have active UC defined as Mayo score of 6 to 12 inclusive, with endoscopic subscore of ≥ 2, a rectal bleeding score of ≥ 1, and a stool frequency score ≥ 1], exclusion criterion #2 [Diagnosis of Crohn's disease or indeterminate colitis or the presence or history of a fistula consistent with Crohn's disease or microscopic colitis or radiation colitis or ischemic colitis], and primary non-responders to 2 or more biologic agents approved for the treatment of UC)

The PP population will be fully identified and documented prior to induction treatment is unblinded. Patients excluded from PP population and the reason for exclusion will be listed. Similar to the ITT population, PP patients in Cohort 1 who received an incorrect investigational drug from that which was randomized will be summarized in the treatment group according to the intended randomization. Patients in Cohort 1 who were randomized with a misreported stratum will be analyzed according to their original [misreported] stratum.

5.3. Safety Population

The Safety population will consist of all patients who received at least 1 dose of investigational drug. This population will be used for all summaries of safety data. Patients randomized to placebo who receive any amount of RPC1063 will be summarized in the RPC1063 1 mg group. Patients randomized to RPC1063 who receive only placebo for all doses will be summarized in the placebo group, otherwise they will be summarized in the RPC1063 1 mg group.

6. STATISTICAL METHODOLOGY FOR EFFICACY

General conventions on data reporting and analysis regarding reporting formats, definitions of Baseline, change, percent change, analysis visit window, methods for handling missing data (e.g., missing Mayo subscore), and methods for handling treatment failures in analysis can be found in Appendix 1.

Unless stated otherwise, the stool frequency and rectal bleeding subscores and associated Mayo composite scores using the 7-day scoring algorithm described in Appendix 3.2 will be used as the primary analysis for the primary and key secondary efficacy endpoints: clinical remission and clinical response.

The 14-day scoring algorithm (A and B) described in Appendix 3.1 will also be used for the analysis of these two endpoints serving as a sensitivity analyses.

For statistical analysis purposes, clinical remission and clinical response will be calculated using the Three-component Mayo definition [redacted]

Cohort 1 (double-blind RPC1063 1 mg once daily or placebo) will be used to formally assess the efficacy endpoints in the IP of the trial.

As Cohort 2 (open-label RPC1063 1 mg once daily) does not have a placebo control, no efficacy endpoints will be formally examined for the IP; however, the efficacy measures shown in Section 4.2.2 will be summarized and described.

As detailed in Section 6.2.1, the primary and 3 key secondary endpoints will be subject to a closed, hierarchical testing procedure in order to control the overall Type I error rate (family-wise error rate or FWER) for testing of these multiple endpoints.

[redacted]

6.1. Analysis of Primary Efficacy Endpoint

The primary efficacy endpoint is the proportion of adult patients in clinical remission (Three-component Mayo definition) at Week 10. The null and alternative hypotheses for the primary efficacy endpoint are:

Null hypothesis H_0: Clinical remission $_{RPC1063}$ = Clinical remission $_{Placebo}$

Alternative hypothesis H_a: Clinical remission $_{RPC1063}$ ≠ Clinical remission $_{Placebo}$

6.1.1. Primary Analysis

The primary analysis of the primary endpoint of proportion of patients in clinical remission (Three-component Mayo definition using 7-day scoring algorithm, see Appendix 3.2) at Week 10 will be carried out in the ITT population using Cochran-Mantel-Haenszel (CMH) test at two-sided 5% level of significance, stratified by corticosteroid use at screening (yes or no), and prior anti-TNF use (yes or no). Results will be expressed as number of patients in remission, remission percentages, weighted difference in remission percentages, odds ratio, and associated two-sided 95% confidence intervals (CIs) and p-values. A bar graph showing the percentages of patients in clinical remission by treatment/cohort with the weighted treatment difference and CMH p-value will be produced. See Appendix 6 for some sample SAS codes.

For patients whose clinical remission (Three-component Mayo definition) at Week 10 cannot be adequately determined (including missing data, discontinuation or lost to follow-up), their clinical remission will be imputed using a non-responder imputation (NRI) approach as described in Appendix 1.3. Patients meeting the TFR criteria prior to the Week 10 visit, their clinical remission at Week 10 will be imputed using NRI as described in Appendix 1.3.

6.1.2. Sensitivity Analyses

Sensitivity analyses will be performed to support the primary analysis (Table 2). These analyses use missing data handling methods including multiple imputation (MI) [redacted] all observed data and tipping point analysis [redacted]

Tipping Point Analysis

A tipping point (TP) sensitivity analysis for the primary endpoint, clinical remission (Three-component Mayo definition) at Week 10, will be performed. The TP method is considered as an alternative strategy for handling missing data [redacted] and the analysis for clinical remission (Three-component Mayo definition) at Week 10 will follow the algorithm for a binary variable provided in the [redacted] This analysis will include all observed data and then vary assumptions about the missing outcomes, by assigning remitters to some missing data and non-remitters to other missing data in the two treatment arms. Tipping point is defined here as the difference in the number of remitters (for clinical remission at Week 10) between the two treatment groups in the missing part at which the study conclusion changes. Unlike the non-responder imputation (NRI) which assigns all missing data to a non-responder for the Week 10 clinical remission, a TP analysis replaces the missing value with either remitter or non-remitter so that the resulting treatment comparison, or level of significance between the two arms, varies. Analysis of each of the imputed dataset will be carried out using the same method for the primary analysis as described in Section 6.1.1. Conducting these sensitivity analyses using the proposed missing data handling methods (NRI, tipping point and multiple imputation), will

support that the comparison of the two treatment groups in clinical remission at Week 10 is robust under different missing data imputation methods.

Multiple Imputation (MI)

A multiple imputation procedure replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. These multiply imputed data sets are then analyzed by using standard procedures for complete data and combining the results from these analyses [redacted]

The following process will be carried out MI for the primary endpoint, clinical remission (3 component Mayo definition) at Week 10. Because clinical remission and response are determined using a 3- component subscores: Rectal Bleeding (RB), Stool Frequency (SF) and Endoscopy subscore, MI will be applied to each of the 3 component subscores followed by applying the definition to determine clinical remission and/or response at Week 10.

The missing data pattern will be examined first. The missing data for RB, SF and Endoscopy subscores is expected to follow either monotone pattern or arbitrary (non-monotone) for each of the three subscores.

- When a subscore (e.g., RB) is considered to follow a monotone missing pattern for a particular individual at a given time point (e.g. Week 5), it implies that all subsequent visits (Weeks 5 and 10) are missing for that individual. For data sets with monotone missing patterns, the variables with missing values can be imputed sequentially, with data from the 2 visits (Weeks 5 and 10) use a logistic regression method (adjusting for the two stratification factors and other baseline covariates deemed appropriate).
- A full conditional method (FCM) will be applied when a subscore (RB or SF) is considered to follow an arbitrary missing pattern.

Combining inferences from imputed datasets involves the following three steps:

- The missing data are filled in m=20 times to generate m complete data sets.
- The m=20 complete data sets after MI are analyzed using the primary analysis method outline in Section 6.1.1.
- The results from the m=20 complete data sets are combined for the inference.

The above MI process results in valid statistical inferences that properly reflect the uncertainty due to missing values; for example, valid confidence intervals for parameters.

The inference of the tipping point analysis will be made based on several factors including missing data pattern, reasons of subject discontinuation/drop-out, results from all observed (non-missing) data and clinical remission results in the published historical studies.

All these analyses (sensitivity and tipping point) will be performed using SAS.

In addition the primary analysis will also be repeated on each of the subgroups listed in Section 6.4. The subgroup analyses will be reported in a table and a forest plot showing the weighted treatment differences with associated 95% CI and p-values will be produced.

These will be considered sensitivity/supportive analyses only and will not be subject to family-wise Type I error control. The analyses below will be calculated using the three-component Mayo definition with 7-day scoring algorithm (Appendix 3.2), unless specified otherwise.

Table 2: Sensitivity Analyses for the Primary Endpoint

Analysis/Endpoint Description	Imputation	Population
Clinical remission – Observed Cases	None	ITT
Clinical remission – Tipping Point Analysis	TPA	ITT
Clinical remission – Multiple Imputation	MI	ITT
Clinical remission	NRI	PP
Clinical remission – 14-day Scoring Algorithm (A, B)	NRI	ITT
[redacted]		
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]

6.2. Analyses of Key Secondary Efficacy Endpoints

The key secondary efficacy endpoints [a priori ordered] are:

- The proportion of adult patients with a clinical response at Week 10
- The proportion of adult patients with endoscopic improvement at Week 10
- The proportion of adult patients with mucosal healing at Week 10

The ordering of the 3 key secondary endpoints reflects the assigned order of the study objectives as outlined in Section 3 and such testing procedure has been shown to control the FWER at alpha of 0.05 [redacted]

6.2.1. Control of Family-wise Type I Error Rate

The primary and 3 key secondary endpoints will be tested in order in a closed, hierarchical testing procedure in order to control the overall Type I error rate for multiple endpoints. If the primary endpoint (Three-component Mayo definition) at Week 10 is statistically significant, the proportion of patients with a clinical response (Three-component Mayo definition) at Week 10 will be tested at the 5% level of significance. If that endpoint is significant, then the proportion of patients with endoscopic improvement at Week 10 will be tested at the 5% level of significance. This testing procedure will continue through each of the 3 key secondary endpoints in the order listed in Section 4.2.2 until an endpoint fails to reach statistical significance, after which all subsequent key secondary endpoints will be considered exploratory. Endpoints listed as other secondary endpoints (Section 4.2.2) will be tested in a nonhierarchical fashion without adjustments for multiplicity.

6.2.2. Primary Analysis of Key Secondary Efficacy Endpoints

The clinical response determined using the 7-day scoring algorithm (Appendix 3.2) and the other two key secondary endpoints (endoscopic improvement and mucosal healing) will be tested using the same type of CMH test as specified for the primary endpoint. For each key secondary endpoint, the same summary and inferential statistics will be reported as specified for the primary analysis of the primary endpoint in Section 6.1.1, including a bar graph showing the percentages by treatment/cohort with the weighted treatment difference, associated two-sided 95% CI and CMH p-value.

Patients without sufficient data to determine clinical response at Week 10 will be analyzed using NRI as described in Appendix 1.3. Patients without central read endoscopy results to determine endoscopic improvement at Week 10 will be analyzed using NRI. Patients without endoscopy or Geboes index results to determine mucosal healing at Week 10 will be analyzed using NRI.

For all key secondary endpoints, patients meeting the TFR prior to the Week 10 efficacy evaluation will be analyzed using NRI as described in Appendix 1.3.

The analysis of each key secondary endpoint will be repeated on each of the subgroups listed in Section 6.4. In addition to reporting the subgroup analyses out in a table, a forest plot showing the weighted treatment differences with associated 95% CI and p-values will be produced.

[redacted]

[redacted]

6.3.1. Main Analysis

Change in Mayo score (complete, [redacted] from Baseline to Week 10 will be analyzed using an analysis of covariance (ANCOVA) model adjusted for corticosteroid use at screening (yes or no), prior anti-TNF use (yes or no), and the Baseline Mayo score, using observed cases. Actual values and changes from Baseline will be descriptively summarized. The adjusted least-squares means and standard errors (SEs) with 95% CI for the difference in mean changes from Baseline and p-values will be reported. The Week 5 visit, when applicable, will also be included in the same table. For example, there will be an analysis for partial Mayo score at Week 5 but not for complete or 9-point Mayo scores.

Proportion of adult patients with histologic remission at Week 10 will each be analyzed and reported using the same stratified CMH test as specified for the primary analysis of the primary endpoint in Section 6.1.1, using NRI described in Appendix 1.3, including a bar graph showing the percentages by treatment/cohort with the weighted treatment difference, associated 95% CI and CMH p-value. [redacted]

Proportion of adult patients with clinical response, remission, or endoscopic improvement at Week 10 in patients who previously received anti-TNF therapy will be analyzed and reported using the same stratified CMH test as specified for the primary analysis of the primary endpoint in Section 6.1.1, except prior anti-TNF use stratification factor will not be used, using NRI described in Appendix 1.3.

[redacted]

[redacted]

6.4. Subgroup Analysis

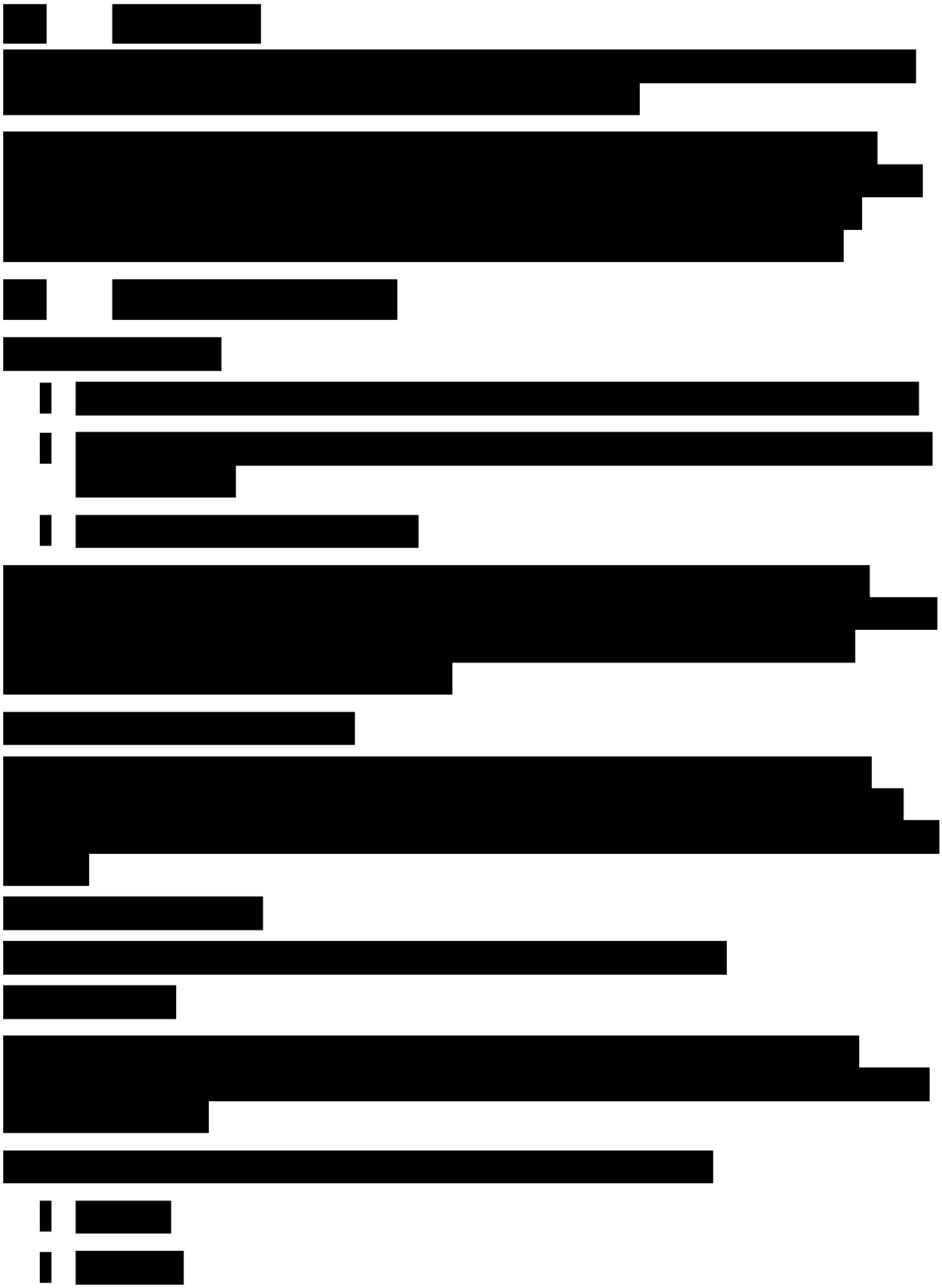
Subgroup analyses will be performed for the endpoints of clinical remission and clinical response only using the three-component Mayo score based on a 7-day scoring algorithm (Appendix 3.2). For each specific subgroup, if there are not enough patients (i.e., <5% ITT population), the corresponding analyses will not be performed. The following are the pre-defined subgroups:

1. Corticosteroid use at screening (yes vs. no)
2. Prior anti-TNF use (yes vs. no)
3. Baseline complete Mayo score (≤ 9 vs. > 9)
4. Extent of colitis (left-sided vs. extensive)
5. Sex (female vs. male)
6. Age at screening (≤ median vs. > median)
7. Baseline fecal calprotectin (≤ 250 vs. > 250 mg/kg)
8. Baseline ALC (≤ 1,500 vs. > 1,500 10^6/L)
9. Years since initial UC diagnosis (≤ 4 vs. > 4 years)
10. Region (North America, Eastern Europe, Western Europe, Asia Pacific)
11. Baseline partial Mayo score (<=median vs. >median)

12. Baseline partial Mayo score (<=7 vs. >7)

13. Baseline endoscopy subscore (2 vs. 3)

Other subgroups such as race (white vs. non-white), weight (≤ median vs. > median), 5-aminosalicylic acid (5-ASA) use at screening (yes vs. no), and corticosteroid and 5-ASA use at screening (yes vs. no), and moderate UC at screening may also be explored.



[redacted]

6.7. Result Blinding

Before unblinding Cohort 1 treatment codes, results of PK, total and differential white blood cell (WBC) count (basophils, eosinophils, lymphocytes, monocytes, and neutrophils), fecal calprotectin, and C-reactive protein beyond first dose will not be disclosed to preserve the blind.

6.8. Interim Analysis

No interim analyses for efficacy are planned during the conduct of the IP.

The data from the IP in Cohort 1 will be cleaned, unblinded, and analyzed after the last Cohort 1 patient has had their last IP visit (I 3 Visit). The analysis of Cohort 1 induction data which is

used for planning for future development will be carried out by designated un-blinded Biostatistics and Programming teams at PPD (a CRO), the same teams already contracted to support the on-going DSMB meetings for this study. A limited number of Sponsor (i.e., Celgene) members who are not part of the study team will have access to the Cohort 1 induction analysis results. The review and conduct of the study will continue in a blinded manner by study team members that are blinded to all study data until all randomized subjects have completed 42 weeks of the Maintenance Phase (Week 52 of overall study duration). The process the Cohort 1 induction analysis will be documented in a charter in a separate document.

The data from Cohort 2 will be cleaned, locked, and analyzed after the last Cohort 2 patient has had their last Cohort 2 visit. The treatment blind for the Maintenance Period will not be broken until the end of the study.

An independent data safety monitoring board (DSMB) will be charged with monitoring accumulating safety data from the trial, as well as general aspects of trial conduct.

The committee will meet periodically during the trial (approximately 4 times each year) to review aggregate analyses by treatment group concerning enrollment, treatment compliance, adherence to follow-up schedule, and safety data from the trial. The DSMB may recommend modifying or stopping the trial early due to safety concerns based on data reviews. The unblinded data will be prepared by an independent team at the CROs.

The blinding plan to assure that all personnel involved in the conduct of the trial remain blinded to the results of interim reviews will be specified in the DSMB Charter.

7. SUMMARY OF PATIENT DISPOSITION

The disposition of patients will be summarized with numbers and percentages by treatment/cohort and in Total Cohort 1 for all randomized subjects (Cohort 1) and all enrolled subjects (Cohort 2). Summaries will include the number and percentage of patients in the following

- Never dosed, Dosed, Completed the IP, and Discontinued from the IP.
- Primary reasons for discontinuation from the IP.

Patients who were randomized to Cohort 1 or enrolled to Cohort 2 will be summarized with number and percentage by region, country, and site for all randomized subjects (Cohort 1) and all enrolled subjects (Cohort 2). Patients who were dosed and who completed the induction period will be summarized with number and percentage by region, country, and site for the ITT population.

Patients in each analysis population as defined in Section 5 will be summarized with number and percentage.

Patients who were screened and who failed screening will be summarized with number and percentage by reason for All Consented population.

Patient disposition will be listed for the ITT population. Patient eligibility will be listed for All Consented population.

8. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Summaries for the demographics, baseline characteristics, medical history, prior medication, and protocol deviations will be presented for the ITT population by treatment/cohort and in Total Cohort 1 unless specified otherwise. Concomitant medications will be presented for the Safety population by treatment/cohort and in Total RPC1063 1 mg. Individual listings will also be provided, including concomitant medical procedures for the Safety population.

8.1. Demographics

Demographic characteristics including age, age category (18-29, 30-39, 40-49, 50-59, 60-69, 70-75, < 65, ≥ 65, ≤ Cohort 1 median, and > Cohort 1 median), sex, race, ethnicity, weight, height, body mass index (BMI), region (North America, Eastern Europe, Western Europe, Asia Pacific, South America, South Africa), country, and tobacco/nicotine usage (never, former, current) will be summarized for the ITT population and the PP population. Demographic characteristics will also be summarized for All Screen Failures, except weight, height, and BMI, which are not collected for screen failure patients. See reporting convention in Appendix 1.1 .

8.2. Ulcerative Colitis Disease Characteristics at Baseline

Ulcerative colitis disease characteristics at baseline including corticosteroid use at screening and prior anti-TNF use as entered in IXRS, actual corticosteroid use at screening and prior anti-TNF use as derived by concomitant medication and prior UC medication eCRFs (corticosteroid and TNF-alpha antagonist, respectively), prior corticosteroid use, prior immunomodulator use, prior biologics use except TNF-alpha antagonist, and prior oral 5-ASA use will be summarized for the ITT population and the PP population. See reporting convention in Appendix 1.1.

The following UC disease history parameters will be summarized in the same table.

- Age at UC symptom onset
- Age at UC diagnosis
- Years since UC symptom onset
- Years since UC diagnosis
- Years since UC diagnosis category (0 to <1, 1 to <2, 2 to <5, 5 to <10, 10 to <15, ≥15, ≤4, >4)
- Extent of colitis (left-sided vs. extensive)
- Complete Mayo score at Baseline using the central endoscopy subscore

- Complete Mayo score at Baseline category using the central endoscopy subscore (≤ 9 vs. > 9)

[redacted]

8.3. Medical History

A summary of medical history, excluding data collected on the UC history and disease status eCRF, will be presented by system organ class (SOC) and preferred term (PT) according to the Medical Dictionary for Regulatory Affairs (MedDRA) version 18.1, or higher, using number and percentage of patients.

[redacted]

9. STUDY TREATMENTS AND EXTENT OF EXPOSURE

All patients will initiate investigational drug via a 7 day dose escalation regimen starting with RPC1063/ozanimod HCl 0.25 mg (equivalent to ozanimod 0.23 mg) or matching placebo (matching placebo for Cohort 1 only) on Days 1 to 4 and RPC1063/ozanimod HCl 0.5 mg (equivalent to ozanimod 0.46 mg) once daily or matching placebo on Days 5 to 7. On Day 8, patients will receive the final dose level (1 mg once daily or matching placebo) for 9 weeks. Patients will record whether they took the daily dose of medication in a diary that will be reviewed periodically by site staff and the Site Monitor.

There will be an individual investigational drug accountability record for each patient and the Investigator should maintain accurate records of the disposition of all investigational drug supplies received during the trial. The number of capsules and date investigational drug supplies are dispensed to the patient and returned by the patient are reported on the Drug Log eCRFs. Overdoses are reported on the Overdose Log eCRF and drug interruptions are reported on the Drug Interruption Log eCRF.

A summary of extent of exposure to IP treatment for the Safety population will include overall treatment duration, dose modifications, and treatment compliance, as defined in Section 9.1, Section 9.2, and Section 9.3. The summary will be presented by treatment/cohort and in Total RPC1063 1 mg. Individual data listings will also be provided.

9.1. Treatment Duration

Additional investigational drug may be dispensed to patients at Week 10 visit (I3) who will continue with maintenance period but could not complete the first maintenance visit (M1) on the same day as I3. To account for this possibility of extended IP treatment, overall IP treatment duration (TDUR_IP) will be defined as:

- the total number of days from the date of first dose to the date of last dose as reported on the End of Study eCRF for patients who do not enter MP, or
- the total number of days from the date of first dose to the date before onsite dosing for MP for patients who enter MP.

If the date of last IP dose is incomplete, see imputation method in Appendix 2 before calculating treatment duration. Overall IP treatment duration will be summarized descriptively.

9.2. Dose Modifications

Dose modifications as reported on the Drug Log, Drug Interruption Log, and Drug Overdose Log eCRFs will be used in summaries. Patients with any IP drug overdose, any IP drug interruption, any IP drug interruption of at least 14 days will be summarized by number and percentage, respectively. The total number of drug overdoses and total number and percentage of AEs as a

result of overdose will be summarized. Duration of drug interruption will be summarized descriptively. The number and percentage of patients will be summarized by duration of drug interruption category (<7 days, 7-<14 days, ≥14 days). The total number of drug interruptions and reasons for interruption will be summarized.

9.3. Treatment Compliance

Treatment compliance for IP will be calculated for each patient by taking into account whether a patient takes all doses of investigational drug as instructed.

For subjects undergoing dose escalation, the total number of capsules expected to be taken by a patient is calculated as follows:

Last dose date – first dose date + 4 (+ 1 to account for the number of days between those dates inclusive and +3 to account for 3 extra capsules taken during dose escalation on Study Days 5-7)

If a patient discontinues trial during or prior to Study Days 5-7 during which 2 capsules per day are taken, then the number of capsules expected will be adjusted accordingly. For example, if discontinuing on Day 5, then +1; on Day 6, then +2; on Day 7 or after, then +3.

The total number of IP capsules actually taken by a patient is defined as the total number of capsules dispensed – total number of capsules returned during IP, as reported on the Drug Log eCRFs. The overall IP treatment compliance (%) will be calculated as the total number of IP capsules taken divided by the total number of IP capsules expected and multiplied by 100. If the bottle is not returned by the patient at a visit, then the subject is considered to have returned zero pill from the unreturned bottle in the treatment compliance calculation.

The total number of IP capsules expected, actually taken and overall IP treatment compliance will be calculated and summarized after taking into account all such visits at which the bottle is not returned.

Additionally, the number and percentage of patients will be summarized by category (<50%, 50% to <60%, 60% to <70%, 70% to <80%, 80% to 100%, and >100%). Two additional categories (<80% and >120%) will also be included.

10. PROTOCOL DEVIATIONS/VIOLATIONS

Deviations and violations from the protocol will be recorded appropriately. Protocol deviations will be classified into, but not necessarily limited to, the following categories:

- Excluded concomitant medication
- Informed consent
- Other/GCP issue
- Randomization procedure
- Selection criteria not met (i.e., violation of inclusion/exclusion criteria)
- Serious adverse event reporting
- Study procedure
- Treatment deviation
- Visit scheduling

Classification of deviations as major protocol deviations/violations and decisions regarding exclusion of patients from the PP will be decided on a case-by-case basis without knowledge of the treatment assigned and before database lock or Cohort 1 unblinding. Major protocol deviations will be presented in a summary table by protocol deviation category. All protocol deviations will be listed. Major protocol deviations that exclude patients from the PP population will be summarized separately.

Besides the listing for patients excluded from the PP population with reason as specified in Section 5.2, a separate listing will be created for patients who have met TFR during IP and thus excluded from PP population. This listing will include the date and study day of their first occurrence of confirmed TFR, and relevant description (e.g., medication name, indication, dose/frequency, Baseline dose/frequency, duration of use, or medical procedure name). This listing will also include the date and study day of their first IP dose, and the date and study day of their W10 endoscopy (if available). This listing will be reviewed, confirmed, and finalized before database lock or Cohort 1 unblinding. See detail about treatment failure rules data review in Appendix 1.3.

11. SAFETY ANALYSIS

All analyses of safety data will be conducted using the Safety population by treatment/cohort and in Total RPC1063 1 mg unless specified otherwise. Individual data listings will also be provided. General conventions on data reporting and analysis regarding reporting formats, definitions of Baseline, change, percent change, analysis visit window, and treatment failure rules can be found in Appendix 1. Analysis visits will be assigned for laboratory, vital signs, ECG, and pulmonary function test (PFT) data as per Appendix 1.3.
No imputation of missing safety data will be performed except in the case of incomplete onset/stop dates for AEs, prior and concomitant medications (Appendix 2).

11.1. Adverse Events

AEs will be monitored during the trial and the data analyzed with respect to incidence within each treatment group as well as severity and potential relationship of the AEs to investigational drug.
A serious adverse event (SAE) is any untoward medical occurrence that at any dose results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing inpatient hospitalization, results in persistent or significant disability/incapacity, is a congenital abnormality/birth defector is an important medical event that may not be immediately life-threatening or result in death or hospitalization but jeopardize the patient or may require intervention to prevent one of the other outcomes listed in the definition above.

A treatment-emergent adverse event (TEAE) is any AE that emerges during treatment having been absent pre-treatment, or worsens relative to the pre-treatment state, excluding those with onset after first MP dose.

An overall summary of TEAEs during IP will include the number and percentage of patients with

- ≥ 1 TEAE
- ≥ 1 Moderate or Severe TEAE
- ≥ 1 Severe TEAE
- ≥ 1 Possible, Probable or Related TEAE
- ≥ 1 Related TEAE
- ≥ 1 Serious TEAE
- ≥ 1 Possible, Probable or Related Serious TEAE
- ≥ 1 Related Serious TEAE

- ≥ 1 TEAE Leading to Discontinuation of Study Drug
- ≥ 1 TEAE Leading to Interruption of Study Drug
- Death
- Death Related to Study Drug

TEAEs will also be summarized and presented by decreasing frequency in the sort keys below in the Cohort 1 RPC1063 1 mg group:

- by system organ class and preferred term: all TEAEs, related TEAEs, severe TEAEs, serious TEAEs, possible, probable or related serious TEAEs, TEAEs leading to interruption of study drug, TEAEs leading to discontinuation of study drug
- by preferred term: all TEAEs, TEAEs with ≥ 2% higher incidence rate in RPC 1 mg than placebo, TEAEs with incidence rate ≥ 5%, TEAEs with incidence rate ≥ 2%
- by system organ class, preferred term, and maximum severity: all TEAEs

Patients with multiple events reported for the same summary level will be counted only once.

11.2. Adverse Events of Special Interest

Potential AEs that may be a consequence of S1P1R modulation will be closely monitored during the trial. These AEs include bradycardia, heart conduction abnormalities (2nd degree and higher AV block), macular edema, malignancy, serious or opportunistic infection, pulmonary effects, and hepatic effects. TEAEs of special interest (AESI) adjudicated by the safety review team per the safety management plan will be summarized by criterion (cardiac, pulmonary, elevated ALT, macular edema, infection, and malignancy) as reported on the Adverse Events eCRF and preferred term. Patients with multiple events reported for the same summary level will be counted only once.

11.3. Deaths

All patient deaths during this trial will be collected and presented in a listing. The listing will include treatment group, period of death, date and study day of last IP dose, date and study day of death, days from last dose, system organ class, preferred term and verbatim of the fatal AE.

11.4. Clinical Laboratory Evaluations

Laboratory assessments will be performed by a central laboratory. All summaries will be based on the standard international (SI) units provided by the central lab. Each patient's hematology, blood chemistry, and urinalysis values will be flagged as "low", "normal", or "high" relative to the normal ranges of the central laboratory. Results from certain laboratory assessments are blinded as per Section 6.7.

Summary statistics of observed values, change and percent changes from Baseline in numeric laboratory parameters will be provided by visit (Baseline, Week 5 [hematology/chemistry/pregnancy only], Week 10, highest value post-baseline during IP, lowest value post-baseline during IP, last available measurement while on treatment during IP). For each laboratory parameter, frequency summaries of categorical shifts from Baseline to each post-Baseline time point (Week 5 [hematology/chemistry/pregnancy only], Week 10, and last available measurement while on treatment during IP) will be provided. For select laboratory parameters (ALC, ANC, WBC, AST, ALT), frequency summaries of patients with abnormalities will be summarized by visit (Baseline, Week 5 [hematology/chemistry/pregnancy only], Week 10, last available measurement while on treatment during IP) and overall.

All laboratory data will be listed, including unscheduled/repeat visits. Unscheduled visits will be included in the calculation of Baseline, highest/lowest/last available on treatment/anytime post-Baseline value.

For each laboratory test, individual patients with any values outside the standard reference range will have all their test results listed.

Hematology

The following laboratory parameters will be included in hematology summary tables: red blood cell (RBC) count, total and differential white blood cell (WBC) count (basophils, eosinophils, lymphocytes, monocytes, and neutrophils), platelet count, hemoglobin (HGB), hematocrit, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), and mean corpuscular hemoglobin concentration (MCHC).

The number and percentage of patients with abnormalities in hematology assessments, defined as ALC < 200 cells/μL, ALC < 500 cells/μL, ANC $< 1{,}000$ cells/μL, and total WBC $> 20{,}000$ cells/μL will be summarized overall and by visit post-Baseline. Patients with at least one abnormality in ALC, ANC or total WBC defined above will be listed.

Blood Chemistry

The following laboratory parameters will be included in the blood chemistry summary tables at the screening visit: sodium, potassium, chloride, calcium, magnesium, phosphate, blood urea nitrogen (BUN), glucose, hemoglobin A1c, albumin, alkaline phosphatase, creatinine, alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma glutamyltransferase (GGT), total bilirubin, conjugated bilirubin, total cholesterol, triglycerides, high-density lipoprotein (HDL), low-density lipoprotein (LDL), and C-reactive protein (CRP). For all other visits the tests will include BUN, hemoglobin A1c, albumin, alkaline phosphatase, creatinine, ALT, AST, GGT, total bilirubin, conjugated bilirubin, and CRP.

The number and percentage of patients with abnormalities in ALT and AST, defined as > 1 x ULN, ≥2 x ULN, ≥3 x ULN, ≥4 x ULN, ≥5 x ULN, ≥8 x ULN, and ≥10 x ULN will be summarized by overall and by visit post-Baseline.

Patients with at least one ALT or AST that is ≥ 3 x ULN will be listed showing all LFTs (ALT, AST, total bilirubin, direct bilirubin, alkaline phosphatase, and GGT) observed across all visits.

Evaluation of drug-induced serious hepatotoxicity (eDISH) plots of maximum ALT/AST values versus Total Bilirubin will be presented for assessment of potential Drug Induced Liver Injury (DILI). Total bilirubin value will be taken from the same visit at which maximum ALT/AST is observed. If maximum ALT/AST is reached at multiple visits, then the highest total bilirubin value from those visits will be plotted.

Urinalysis

The following laboratory parameters will be included for urinalysis: leukocytes, specific gravity, bilirubin, blood, glucose, ketones, pH, protein, and urobilinogen.

Serum pregnancy and serology testing

Serum β-human chorionic gonadotropin (hCG) will be collected at Screening for females of childbearing potential. Serology testing will also be collected at Screening and will include immunoglobulin G (IgG) antibodies to varicella-zoster virus (VZV), human immunodeficiency virus (HIV), immunoglobulin M (IgM) antibodies to hepatitis A virus (HAV), surface antigen of the hepatitis B virus (HBsAG), anti-hepatitis B core total antibodies (anti-HBc total), hepatitis B virus deoxyribonucleic acid (HBV DNA), and anti-hepatitis C virus (HCV) IgG/IgM.

Serum pregnancy results and serology testing results at Screening will be listed only.

11.5. Vital Sign Measurements

Systolic and diastolic blood pressure (SBP and DBP) and pulse (HR) will be assessed in a sitting position and collected at each visit. On Day 1 of treatment for the IP, careful cardiac monitoring of the patients is required. Resting pulse and blood pressure in the sitting position will be measured before the first dose of investigational drug and every hour for at least 6 hours thereafter. For the hourly measurements after investigational drug administration, HR and sitting blood pressure will be measured once and recorded in the eCRF. The first dose monitoring strategy should also be repeated at Day 5 or at Day 8 if any cardiac safety issues were observed on the previous Day of dose escalation.

Summary statistics of actual values and changes from Baseline in vital signs will be provided by visit (Baseline, Week 5, Week 10, highest value post-baseline during IP, lowest value post-baseline during IP, last available measurement while on treatment during IP). On Day 1, hourly

changes from predose up to hour 6 in blood pressures and HRs will be summarized for all patients.

Patient's lowest heart rate during 6-hour cardiac monitoring on Day 1 will be summarized by number and percentage by category: ≥65, 60-64, 55-59, 50-54, 45-49, 40-44, 35-39, <35 bpm, as well as <40 bpm and <45 bpm.

The number and percentage of patients requiring extended cardiac monitoring beyond hour 6 on Day 1 or requiring monitoring on Study Days 5 and/or 8 will be summarized by Day: Day 1, Day 5, and Day 8. A listing of these patients will be presented that shows the hourly change from the pre-dose assessment at each of the available Baseline, Study Day 5, and Study Day 8 visits by patient for HR and blood pressures.

The number and percentage of patients with clinically relevant abnormalities defined below will be summarized overall and by visit (Week 5, Week 10, last available while on treatment during IP) and overall. The denominator will be the number of patients with observed data at both Baseline and respective post-Baseline visit, unless specified otherwise.

- SBP:
    - >140 mmHg post-Baseline
    - >160 mmHg post-Baseline
    - >180 mmHg post-Baseline or an increase from baseline of > 40 mmHg
        - >180 mmHg post-Baseline (denominator will be the number of patients with Baseline SBP ≤180)
        - an increase from Baseline of > 40 mmHg
    - <90 mmHg post-Baseline or a decrease from Baseline of > 30 mmHg
        - <90 mmHg post-Baseline
        - <90 mmHg post-Baseline (denominator will be the number of patients with Baseline SBP ≥90)
        - a decrease from Baseline of > 30 mmHg
- DBP:
    - >90 mmHg post-Baseline
    - >100 mmHg post-Baseline
    - >105 mmHg post-Baseline or an increase from Baseline of > 30 mmHg

        - >105 mmHg post-Baseline (denominator will be the number of patients with Baseline DBP ≤105)
        - an increase from Baseline of > 30 mmHg
    - <50 mmHg post-Baseline or a decrease from Baseline of > 30 mmHg
        - <50 mmHg post-Baseline
        - <50 mmHg post-Baseline (denominator will be the number of patients with Baseline SBP ≥50)
        - a decrease from baseline of > 30 mmHg
- HR:
    - >100 bpm post-Baseline
    - >120 bpm post-Baseline or an increase from baseline of > 20 bpm
        - >120 bpm post-Baseline (denominator will be the number of patients with Baseline HR ≤120)
        - an increase from Baseline of > 20 bpm
    - <60 bpm post-Baseline
    - <45 bpm post-Baseline or a decrease from Baseline of > 20 bpm
        - <45 bpm post-Baseline (denominator will be the number of patients with Baseline HR ≥45)
        - a decrease from Baseline of > 20 bpm

The number and percentage of patients with clinically relevant abnormalities during cardiac monitoring, including those beyond hour 6 on Day 1, will be summarized overall and by hour during pre-defined 6-hour monitoring window on Day 1, as well as overall during cardiac monitoring (Days 1/5/8),. The denominator will be the number of patients with observed data at both Baseline and respective post-Baseline timepoint.

Discharge Status eCRF data after receiving extended cardiac monitoring after 6 hours will be summarized by number and percentage of patients. All data reported on the Discharge Status eCRF will be listed, including the reason for extended monitoring.

Box plots will be provided to display the distribution of vital signs results by timepoint on Day 1 up to hour 6, or by visit. In addition, line plots of means and SEs will be created to display the

actual values and change from pre-dose by timepoint on Day 1 up to hour 6. Line plots of means and SEs will also be created to display the actual values and change from Baseline by visit.

11.6. Physical Examination

Physical examination results will be listed only. Patients with at least one dermatological abnormality will be listed.

11.8. Electrocardiograms

Electrocardiograms will be performed before and 6 hours after the first dose of investigational drug administration for all patients on Induction Day 1 while the patient is in the clinic. The 6 hour postdose ECG will be evaluated by the treating physician, with input if needed from a local cardiologist or a central reader to confirm if extended monitoring is required. Additional ECG monitoring will be performed on Days 5 and 8 if cardiac issues are identified on the prior day of dose escalation.

Summary tables including actual values and change from Baseline will be presented for ECG results by visit (Baseline, Day 1 Hour 6, Week 10, last available measurement while on treatment during IP). An ECG outlier analysis will be performed by visit (Baseline, Day 1 Hour 6, Week 10, last available measurement while on treatment during IP) and overall showing the proportion of patients with QT > 450 milliseconds (ms), QT > 480 ms, QT > 500 ms, QTcF > 450 ms, QTcF > 480 ms, QTcF > 500 ms, change from Baseline in QT of >30 ms, change from Baseline in QT of >60 ms, change from Baseline in QTcF of >30 ms, and change from Baseline in QTcF of >60 ms. The overall interpretation of the ECG results (normal; abnormal, probably non-significant; and abnormal, possibly significant) will be summarized in the same table. The denominator will be the number of patients with result (actual value or change) at the respective visit.

11.9. Pulmonary Function Tests

The following pulmonary function tests (PFT) are performed at screening and Week 10: forced expiratory volume in 1 second (FEV_1), forced vital capacity (FVC), percent of predicted FEV_1, percent of predicted FVC, and if locally available, diffusion capacity of the lung for carbon monoxide (DLCO). Summary tables including actual values, change, and percent change from Baseline will be presented for PFT results by visit (Baseline, Week 10, last available

measurement while on treatment during IP). DLCO results in domestic unit (mL/min/mmHg) will be converted to SI unit (mmol/min/kPa) prior to analysis using the following conversion factor:

- DLCO in SI unit = (DLCO in domestic unit) / 2.986

Patients with at least one FEV_1 or FVC < 70% predicted will be listed showing all PFTs observed across all visits. Data collected on the DLCO eCRF will be listed separately.

[redacted]

[redacted]