STATISTICAL ANALYSIS PLAN
Protocol PTG-100-02

A PHASE 2B RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, PARALLEL ADAPTIVE 2-STAGE, MULTI-CENTRE STUDY TO EVALUATE THE SAFETY AND EFFICACY OF ORAL PTG-100 INDUCTION IN SUBJECTS WITH MODERATE TO SEVERE ACTIVE ULCERATIVE COLITIS

Protocol Number:	PTG-100-02
(Version Date)	08 SEPT 2016
Name of Test Drug:	PTG-100
Phase:	2b
Methodology:	Randomized, Double-Blind, Placebo-Controlled, Parallel Adaptive 2-Stage Multi-Center Trial
Sponsor:	Protagonist Therapeutics, Inc. 521 Cottonwood Drive, Suite 100 Milpitas, CA 95035
Sponsor Representative:	Richard S Shames Chief Medical Officer
Document Date:	08 SEPT 2016
Document Version:	1.0

Confidentiality

This document is confidential and proprietary property of Protagonist and to be used only as authorized by Protagonist. No part is to be reproduced, disclosed to others, or quoted without prior written authorization from Protagonist.

SIGNATURE PAGE

Protocol Title: A Phase 2b Randomized, Double-Blind, Placebo-Controlled, Parallel Adaptive 2-Stage, Multi-Centre Study To Evaluate The Safety And Efficacy Of Oral PTG-100 Induction In Subjects With Moderate To Severe Active Ulcerative Colitis

Sponsor: Protagonist Therapeutics, Inc.
521 Cottonwood Drive, Suite 100
Milpitas, CA 95035

Protocol Number: PTG-100-02

Document Date/Version: 08 SEPT 2016 V0.1

Cytel, Inc. Author:
Adam Hamm, PhD
Cytel, Inc.
675 Massachusetts Avenue
Cambridge, MA 02139

Signature: _____

Date: 8 SEP 2016

Sponsor Approval

By signing this document, I acknowledge that I have read the document and approve of the planned statistical analyses described herein. I agree that the planned statistical analyses are appropriate for this study, are in accordance with the study objectives, and are consistent with the statistical methodology described in the protocol, clinical development plan, and all applicable regulatory guidances and guidelines.

I have discussed any questions I have regarding the contents of this document with the biostatistical author.

I also understand that any subsequent changes to the planned statistical analyses, as described herein, may have a regulatory impact and/or result in timeline adjustments. All changes to the planned analyses will be described in the clinical study report (CSR).

Sponsor Signatory:
Richard S. Shames, MD
Chief Medical Officer
Protagonist Therapeutics, Inc.
521 Cottonwood Drive, Suite 100
Milpitas, CA 95035

Signature: _____

Date: 8 SEPTEMBER 2016

TABLE OF CONTENTS

Section **Page**

1. INTRODUCTION AND OBJECTIVES OF ANALYSIS 6
1.1. Introduction 6
1.2. Objectives of Statistical Analysis 6
2. STUDY DESIGN 7
2.1. Synopsis of Study Design 7
2.2. Randomization Methodology 7
2.3. Stopping Rules and Unblinding 8
2.4. Study Procedures 8
2.5. Efficacy, Pharmacokinetic, and Safety Variables 12
2.5.1. Efficacy Variables 12
2.5.2. Pharmacokinetic Variables 13
2.5.3. Pharmacodynamic and Immunogenicity Variables 13
2.5.4. Safety Variables 13
3. SUBJECT POPULATIONS 14
3.1. Population Definitions 14
3.2. Protocol Violations 14
4. STATISTICAL METHODS 15
4.1. Sample Size Justification 15
4.2. General Statistical Methods and Data Handling 15
4.2.1. General Methods 15
4.2.2. Computing Environment 16
4.2.3. Methods of Pooling Data 16
4.2.4. Adjustments for Covariates 16
4.2.5. Multiple Comparisons/Multiplicity 16
4.2.6. Subpopulations 16
4.2.7. Withdrawals, Dropouts, Loss to Follow-up 16
4.2.8. Missing, Unused, and Spurious Data 17
4.2.9. Visit Windows 18
4.3. Interim Analyses 20
4.4. Subject Disposition 21

Section **Page**

4.5. Demographic and Baseline Characteristics 22
4.6. Efficacy Evaluations 22
4.6.1. Primary Efficacy Analysis 22
4.6.2. Secondary Efficacy Analyses 24
4.6.3. Exploratory Analyses 25
4.7. Pharmacokinetic and Pharmacodynamic Evaluations 25
4.7.1. Pharmacokinetic Evaluations 25
4.7.2. Pharmacodynamic Evaluations 26
4.8. Safety Analyses 27
4.8.1. Adverse Events 27
4.8.2. Laboratory Data 28
4.8.3. Vital Signs and Physical Examinations 28
4.8.4. Electrocardiogram 28
4.8.5. Concomitant Medications 29
5. CHANGES TO PLANNED ANALYSES 30
6. References 31
7. CLINICAL STUDY REPORT APPENDICES 32
7.1. Statistical Tables to be Generated 32
7.2. Data Listings to be Generated 37

ABBREVIATIONS

Abbreviation	Definition
ADRC	Adaptive design review committee
AE	Adverse events
AESI	Adverse events of special interest
AUC	Area under the concentration versus time curve
AUC_{0-t}	Area under the concentration-time curve from time 0 to the time of the last sample collection
C_{max}	Maximum plasma concentration
C_{trough}	trough (predose) concentration
CRF	Case Report Form
CRO	Contract Research Organization
CSR	Clinical study report
DMC	Data monitoring committee
ECG	Electrocardiogram
FDA	Food and Drug Administration
IA	Interim Analysis
ICH	International Conference on Harmonisation
IXRS	Interactive web/voice response system
MAR	Missing at Random
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model Repeated Measures
MNAR	Missing Not at Random
PK	Pharmacokinetic
RE	Receptor expression
%RO	Receptor occupancy
SAE	Serious Adverse Event
SAP	Statistical analysis plan
T_{max}	Time to maximum plasma concentration
TNF-α	Tumor necrosis factor - alpha

1. INTRODUCTION AND OBJECTIVES OF ANALYSIS

1.1. Introduction

This Statistical Analysis Plan (SAP) has been developed for the final analysis for protocol PTG-100-02, **a Phase 2b Randomized, Double-Blind, Placebo-Controlled, Parallel Adaptive 2-Stage, Multi-Centre Study To Evaluate The Safety And Efficacy Of Oral PTG-100 Induction In Subjects With Moderate To Severe Active Ulcerative Colitis.**

This SAP outlines the safety and efficacy analyses, including tables, listings and graphical presentations to be included in the Clinical Study Report (CSR). Additional analyses may be requested by the Sponsor after review of the pre-planned analyses.

1.2. Objectives of Statistical Analysis

This statistical analysis plan (SAP) is designed to outline the methods to be used in the analysis of study data in order to answer the study objective(s). Populations for analysis, data handling rules, statistical methods, and formats for data presentation are provided. The statistical analyses and summary tabulations described in this SAP will provide the basis for the results sections of the clinical study report (CSR) for this trial.

This SAP also will outline any differences in the currently planned analytical objectives relative to those planned in the study protocol.

2. STUDY DESIGN

2.1. Synopsis of Study Design

This will be a randomized, double-blind, placebo-controlled, multi-center, parallel adaptive 2-stage design study. It is anticipated that approximately 240 male and non-gravid female subjects aged 18 to 80 years with moderate to severe active ulcerative colitis will be randomized. (If only two arms are selected for continuation following the interim analysis, then up to approximately 210 subjects will be randomized). Randomization will be stratified by prior treatment with Tumor necrosis factor-alpha (TNF-α) antagonist use.

Stage 1 is a four-arm randomized evaluation of three dose regimens of PTG-100 (150 mg, 300 mg, and 900 mg) compared with placebo. After approximately 60-80 subjects have completed the full 12 weeks of dosing, an interim analysis will be performed leading to selection of dose(s) that will be discontinued in Stage 2. Subjects will continue to be randomized across the 4 arms during the interim analysis until dose(s) are selected for Stage 2 of the trial.

Stage 2 is a randomized evaluation of the dose(s) of PTG-100 selected based on the interim analysis data compared with placebo. Since the trial provides for a 12-week treatment period, it is expected that an additional approximately 60 subjects would be randomized in equal proportions to the 4 treatment groups while the first approximately 60-80 subjects' 12-week data are accumulated and analyzed. These Stage 2 subjects will be referred to as the "overrun", and will be included in all final analyses, but not in the interim analysis. Subjects in the continuing arms who are randomized after Stage 1 and before the end of the interim analysis period will be included in all analyses as Stage 2 subjects. Hence, in Stage 2 the remaining subjects would be randomized in equal proportions to the selected dose(s) and placebo. This would yield approximately 70-80 subjects per treatment group for the selected doses. Final analysis will combine all observed data from both Stages 1 and 2.

Stage 2 Modification Post-Interim Analysis: It is possible that a new treatment arm (600 mg) not tested in Stage 1 could be added to the trial in Stage 2. It is to be considered as exploratory only, and the data collected on subjects randomized to this arm will not be included in any hypothesis testing for overall trial results. The data from this arm will be compared only to the data of subjects randomized to the placebo arm, for all trial endpoints.

2.2. Randomization Methodology

Subjects are randomized 1:1:1:1 by interactive web/voice response system (IXRS) in Stage 1. Subjects will be stratified by prior treatment with TNF-α antagonist use (Yes/ No). Subject enrollment will occur in 2 stages relative to the interim analysis. An independent Data Monitoring Committee (DMC) will monitor safety throughout the study period and may be unblinded as necessary. Limited personnel required for conducting the analyses may also be unblinded to allow necessary safety oversight. The DMC will also be involved in the interim analysis. The purpose of the interim analysis will be to conduct a futility analysis and to identify which treatment arms provide optimal data in order to select those (one or two) PTG-100 dose levels for continual enrollment as the most informative and effective dose arms. Dose selection

will further be made in consideration of safety data and sound clinical judgment, and will be made by the Adaptive Design Review Committee (ADRC), which will be comprised of the DMC and a representative of the Sponsor (not involved in the direct conduct of the study).

After the unblinded interim analysis by the ADRC, additional subjects will be randomized equally (1:1 or 1:1:1 as appropriate) to the selected doses of PTG-100 and placebo with stratification maintained in Stage 2.

2.3. Stopping Rules and Unblinding

Conditional power will also be computed at the interim analysis. It is the probability that the trial will yield a statistically significant difference for clinical remission at the higher chosen dose for Stage 2 if the observed difference at the Interim Analysis (IA) is the TRUE underlying difference and the trial is continued to completion. If this probability is less than 10%, the decision could be made by the ADRC to stop the trial for futility after review of safety data and other efficacy and/or exploratory endpoints.

This will be a double-blind, placebo-controlled study. As such, the sponsor (other than the designated representative of the ADRC), CRO, medical monitors, investigator, site pharmacist (or designee), site personnel involved in study conduct, and subjects will remain blinded to study medication assignment.

An independent Data Monitoring Committee (DMC) will monitor safety throughout the study period and may be unblinded as necessary. Limited personnel required for conducting the analyses may also be unblinded to allow necessary safety oversight. The DMC will also be involved in the interim analysis. Limited personnel required for conducting the interim analysis will also be unblinded. Unblinded personnel required for the interim analysis will be defined in the DMC charter. The sponsor representative serving on the ADRC will not be involved in the periodic safety monitoring performed by the DMC.

Unblinding of individual subject treatment assignments may be permitted in the event of an emergency situation requiring knowledge of treatment assignment; however, subject assignment should not be routinely unblinded. Unblinding will be performed via the IXRS system from Almac (Craigavon, Northern Ireland). This randomization information may be opened and the randomization for each individual subject made available to the Investigator (or designee) only in the event of a medical emergency or an adverse event (AE) that necessitates identification of the clinical trial material for the welfare of that subject. Except in a medical emergency, the Investigator (or designee) and blinded clinical site staff will remain blinded during the conduct of the study and until such time that all discrepancies in the clinical database are resolved (i.e., at the database lock). The date/initials and reason for the Investigator (or designee) removing the study blind will be documented. Unblinding for any other reason will be considered a protocol violation.

2.4. Study Procedures

The schedule of assessments, as outlined in the study protocol, is provided in Table 1.

Table 1 Schedule of Assessments

	Screening	Random-ization Period	Treatment								Follow up	
	Days -35 to -7	**Day -7 to Day -1**	**Day 0 (pre-dose)[1]**	**Day 0 (dosing, post-dose)**	**Day 14 ±3 days**	**Day 28 ±3 days**	**Day 42 ±3 days**	**Day 56 ±3 days**	**Day 70[2] ±3 days**	**Day 84 ±3 days**	**Day 112 ±7 days**	**Early Termi-nation[3]**
Informed Consent	X											
Demographics	X											
Medical History	X											
Prior and Concomitant Medications[4]	X		X	X	X	X	X	X	X	X	X	X
Inclusion/Exclusion	X											
Height	X											
Weight	X		X		X	X	X	X		X	X	X
Physical Examination[5]	X		X	X	X	X	X	X		X	X	X
Vital Signs[6]	X		X	X[7]	X	X	X	X		X	X	X
12-lead ECG[6]	X		X[8]		X			X		X	X[9]	X
Laboratory: Haematology, Coagulation, Serum Chemistry[10]	X		X		X	X		X		X	X	X
Stool *C.diff*, Culture, and Ova and Parasites	X											
Urinalysis (dipstick and microscopy if abnormal)	X		X		X	X		X		X	X	X
PML Neurological Assessment[11]	X		X		X	X	X	X	X	X	X	X

	Screening	Random-ization Period	Treatment								Follow up	
	Days -35 to -7	**Day -7 to Day -1**	**Day 0 (pre-dose)[1]**	**Day 0 (dosing, post-dose)**	**Day 14 ±3 days**	**Day 28 ±3 days**	**Day 42 ±3 days**	**Day 56 ±3 days**	**Day 70[2] ±3 days**	**Day 84 ±3 days**	**Day 112 ±7 days**	**Early Termi-nation[3]**
Hepatitis B & C; TB Test	X											
Serum Pregnancy (WOCBP)	X									X	X	X
Sigmoidoscopy	X[12]									X		X
Colonic Biopsy	X[12]									X		X
Mayo Score[13]	X[12]				X	X	X	X	X	X	X	X
JCV Antibody Testing[14]	X[14]											
Confirmation of Eligibility			X									
Randomisation[15]		X										
Adverse Events			X	X	X	X	X	X	X	X	X	X
IBDQ			X							X		X
Faecal Calprotectin			X				X			X	X	X
Urine Pregnancy (WOCBP)			X		X	X	X	X				
PK Blood Collection[16]			X	X[17]	X	X	X	X		X[17]		X[17]
PD Blood Collection[16]			X	X[17]	X	X	X	X		X[17]		X[17]
ADA Testing			X			X		X		X	X	X
Exploratory Immunologic Assessment Blood Sample[18]			X									
IP Administration[19]				X	X	X	X	X	X	X		

ADA = anti-drug antibodies; AE = adverse event; ECG = electrocardiogram; IBDQ = Inflammatory Bowel Disease Questionnaire; IP = investigational product; JCV= John Cunningham virus; PD = pharmacodynamics; PK = pharmacokinetics; PML = progressive multifocal leukoencephalopathy; TB = tuberculosis; WOCBP = women of childbearing potential.

[1] All predose procedures should be completed prior to dosing.

[2] Day 70 AE monitoring will be obtained via phone call and data to calculate rectal bleeding and stool frequency subscores will be obtained.

[3] Procedures to be performed within 10 days of Early Termination.

[4] All prior medications taken within 3 months before screening should be recorded.

[5] Complete physical examinations should be conducted at Screening, Day 84, Day 112, and/or Early Termination. Directed physical examinations should be conducted at other time points. The Day 0 postdose examination will occur 4 hours postdose.

[6] Participants should be rested in a supine or semi-recumbent position for ≥ 3 minutes prior to recording of ECGs and vital signs. The ECGs and vital signs should be assessed prior to blood draws. Temperature by oral, axillary, or tympanic route may be permitted provided the same route is used for the participant throughout the study.

[7] Vital signs will be done at 1, 2, and 4 hours ± 20 minutes postdose following the first dose of study drug.

[8] Predose (baseline) ECGs should be repeated 3 times with 1- to 2-minute intervals between ECG readings, all within 2 hours prior to dosing.

[9] ECG only to be performed if previously abnormal.

[10] Laboratory samples collected at Screening and on Day 0 (predose), Day 28, Day 84 (or Early Termination), and Day 112 MUST be after the participant has FASTED from all food and drink except water for ≥ 8 hours (except for any dietary requirements that are part of sigmoidoscopy prep). Chemistry Panel at these fasting visits will include cholesterol and glucose analytes. All other non-fasting visits will have Chemistry Panel completed without cholesterol and glucose analytes. See Appendix B for list of analytes assessed.

[11] PML Assessment Questionnaire (see Appendix D). If the Screening PML assessment was performed within 14 days of Day 0, it will not need to be repeated on Day 0.

[12] Sigmoidoscopy, biopsy, and a baseline complete Mayo Score are to be completed within 14 days of randomisation (as close to randomisation as possible) after confirmation of eligibility from other screening tests. Eligibility/ baseline complete Mayo Score should be done before the day of sigmoidoscopy. Sigmoidoscopy should be the last procedure before randomisation. Subjects must meet eligibility criteria of Mayo Score 6 to 12 and endoscopy subscore of ≥ 2 in order to be randomised.

[13] Complete Mayo Score (including all 4 subscores) to be assessed at Screening, Day 84, and Early Termination (if possible); partial Mayo Score (without endoscopy subscore) to be assessed at all other time points except Day 70 where only rectal bleeding and stool frequency subscores will be reported via phone.

[14] JCV antibody testing must be performed within 90 days prior to dosing. Testing will be performed at the time of the pretreatment sigmoidoscopy if no prior results (within the 90-day timeframe) are available, and if the subject is eligible for study enrolment. Results will not determine eligibility for the trial.

[15]Randomisation will occur after all subject eligibility has been confirmed in the screening period. Randomisation in the interactive web/voice response system must be completed in advance of Day 0. Dosing should occur within 7 days of randomisation.

[16] Baseline (Day 0 predose) PK and PD blood samples will be taken within 2 hours prior to dosing. PD samples to be taken at selected sites only.

[17] PK samples are to be collected at 1, 2, and 4 hours ± 20 minutes postdose on Day 0 and Day 84 or Early Termination (if possible) only. PD samples to be collected 4 hours ± 20 minutes postdose on Day 0 and Day 84 or Early Termination (if possible) only. For every visit other than Day 0, a single PK and PD sample will be obtained within 1 hour prior to taking the regularly scheduled dose.

[18] A separate blood sample will be collected at selected sites (Day 0 predose only) for immunologic assessment.

[19] IP will be administered once daily for 84 days. The first drug will be administered at the study centre and subjects will be observed for at least 4 hours. On visits to study site, subjects should take the daily dose of study medication after completion of all blood collections.

2.5. Efficacy, Pharmacokinetic, and Safety Variables

2.5.1. Efficacy Variables

The primary efficacy endpoint for this study is the proportion of subjects receiving PTG-100 with clinical remission at Week 12 compared with placebo.

Clinical remission is defined as follows using the Mayo subscores of stool frequency, rectal bleeding, and endoscopy:

1. Stool frequency subscore of 0 or 1 with a pre-specified change of 1 or more from baseline
2. Rectal bleeding subscore of 0
3. Endoscopy subscore of 0 or 1 (modified so that a score of 1 does not include friability)

Secondary endpoints, all based on comparison of individual PTG-100 dose levels with placebo, include:

1. Proportion of subjects with clinical response at Week 12 (Day 84) (defined as at least 1 point and 30% reduction from baseline in rectal bleeding and stool frequency subscores)
2. Proportion of subjects with endoscopic response at Week 12 (Day 84) (defined as an endoscopic subscore of 0 or 1)
3. Mean change in endoscopy subscore from baseline to Week 12 (Day 84)
4. Mean change in rectal bleeding subscore from baseline to Weeks 2, 4, 6, 8, 10, 12, and 16 (Days 14, 28, 42, 56, 70, 84, and 112)
5. Mean change in stool frequency subscore from baseline to Weeks 2, 4, 6, 8, 10, 12, and 16 (Days 14, 28, 42, 56, 70, 84, and 112)
6. Proportion of subjects with endoscopic remission at Week 12 (defined as an endoscopic subscore of 0)
7. Mean change in Complete Mayo Score (including all four subscores) from baseline to Week 12 (Day 84)
8. Mean change in Partial Mayo Score (excluding endoscopy subscore) from baseline to Weeks 2, 4, 6, 8, 10, 12, and 16 (Days 14, 28, 42, 56, 70, 84, and 112)
9. Mean change in fecal calprotectin levels from baseline to Weeks 6 (Day 42) and 12 (Day 84)
10. Mean change in IBDQ score from baseline to Week 12 (Day 84)

Exploratory endpoints include:

1. Mean change in histological score from baseline to Week 12
2. Effects of ADA on PK, safety, and efficacy in subjects with positive ADA

2.5.2. Pharmacokinetic Variables

Blood samples for determination of PTG-100 pharmacokinetics were to be collected as outlined in Table 1. Maximum concentration (C_{max}), trough (predose) concentration (C_{trough}) and the time of C_{max} (T_{max}) will be taken directly from the observed data. The area under the concentration-time curve (AUC_{0-t}) from time 0 to the time of the last sample collection will be calculated using the linear trapezoidal rule.

2.5.3. Pharmacodynamic and Immunogenicity Variables

Quantitation of percent α4β7 receptor occupancy (%RO) and α4β7 receptor expression (RE) on peripheral blood T cells will be evaluated as secondary endpoints.

ADA will be categorized as screening, confirmatory, or neutralizing. The proportion of subjects developing neutralizing anti-drug antibodies (ADA) by Week 12 (Day 84) and Week 16 (Day 112) will be evaluated.

Exploratory endpoints will include:

1. Analysis of peripheral blood lymphocytes in subjects on individual PTG-100 dose levels compared to placebo
2. Number of β7 positive cells in colonic biopsies as assessed by immunohistochemistry (IHC) in subjects on individual PTG-100 dose levels compared to placebo
3. Changes in immunologic/PD biomarkers in the target population

2.5.4. Safety Variables

Safety assessments performed during the study included physical examinations, measurement of vital signs, 12-lead electrocardiograms (ECGs), clinical laboratory evaluations including hematology, serum chemistry, and urinalysis, concomitant medications, and monitoring of AEs.

3. SUBJECT POPULATIONS

3.1. Population Definitions

The following subject populations will be evaluated and used for presentation and analysis of the data:

Safety Population: all randomized subjects who received any amount of study drug.

Full Analysis Set (FAS): all randomized subjects who received any amount of study drug and who have results from baseline and from ≥1 post-baseline assessment will be included in the PD and efficacy analysis.

Per Protocol (PP) Population: all subjects in the FAS without any major protocol violations

Pharmacokinetic (PK) Population: all randomized subjects who receive a Day 0 dose of study drug and who have sufficient PK data for analysis (meeting the required minimum blood volume collected along with the sampling schedule, and sampling time, and without vomiting)

Subjects in the all populations will be included in the respective treatment groups on an as-treated basis.

The safety population will be used for safety analyses, demographic and other baseline subject characteristics, and the FAS and PP populations will be used for the primary, secondary, exploratory, PD, and immunogenicity endpoints. The FAS population will also be used for the Disposition by Study Visit tabulation. The PK population will be used for PK summaries.

If it is determined that a significant number of subjects were not treated as randomized, additional analyses on the subjects analyzed as randomized may be performed.

Of note, per the adaptive study design, the overrun subjects randomized in Stage 2 will be included in all analyses where data is available.

As noted in Section 2.1, subjects randomized during Stage 2 to a dose not studied in Stage 1 will be included in final analyses.

3.2. Protocol Violations

At the discretion of the sponsor, major protocol violations as determined by a review of the data prior to unblinding of the study results and the conduct of statistical analyses may result in the removal of a subject's data from the Per Protocol Population. The sponsor, or designee, will be responsible for producing the final protocol violation file (formatted as a Microsoft Excel file), in collaboration with Cytel and the data monitoring group as applicable; this file will include a description of the protocol violation, and clearly identify whether or not this violation warrants exclusion from the Per Protocol Population. This file will be finalized prior to hard database lock.

All protocol violations will be presented in the data listings.

4. STATISTICAL METHODS

4.1. Sample Size Justification

Approximately 240 male and non-gravid female subjects aged 18-80 with moderate to severe active UC will be randomized.

Prior studies[1-7] suggest treatment effects of ~21 percentage point differences from placebo clinical remission rate of ~10%. Those values were assumed for purposes of design of this 12-week treatment trial.

Table 2 below shows power for representative scenarios of potential TRUE underlying dose-response relationships, as well as that for no drug effect to assess type 1 error (calculations via Cytel's EAST COMPASS adaptive dose-finding software based on 1000 simulations of each design, with each simulated trial analyzed as described in Section 4.6).

Table 2 Power and Type 1 Error for Various Response Rates

TRUE underlying response rates				power (alpha=0.025, 1-sided)			
placebo	150 mg	300 mg	900 mg	N=260	N=240	N=220	N=200
0.1	0.2	0.31	0.31	93%	91%	90%	84%
0.1	0.17	0.24	0.31	83%	82%	78%	75%
0.1	0.1	0.1	0.1	3.4%	2.2%	1.7%	2.1%

4.2. General Statistical Methods and Data Handling

4.2.1. General Methods

All output will be incorporated into Microsoft Excel or Word files, sorted and labeled according to The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) recommendations, and formatted to the appropriate page size(s).

Tabulations will be produced for appropriate demographic, baseline, efficacy, pharmacokinetic, pharmacodynamic and safety parameters. For categorical variables, summary tabulations of the number and percentage within each category (with a category for missing data) of the parameter will be presented. For continuous variables, the mean, median, standard deviation, minimum and maximum values will be presented.

Formal statistical hypothesis testing will be performed on the primary efficacy endpoint with all tests conducted at the 1-sided, 0.025 level of significance. Summary statistics will be presented, as well as confidence intervals on selected parameters, as described in the sections below.

4.2.2. Computing Environment

All descriptive statistical analyses will be performed using SAS statistical software (Version 9.2), unless otherwise noted. Medical History and adverse events will be coding using the most current MedDRA version available. Concomitant medications will be coded using the most current World Health Organization (WHO) Drug version available and WHODRUG DD B Format.

4.2.3. Methods of Pooling Data

Data will be pooled across all study sites for analyses. Homogeneity of the treatment effect across sites will be evaluated using a logistic regression model with response rate as the dependent variable and terms for treatment, site, and treatment by site interaction as fixed effects. Sites with small enrollment numbers will be pooled with the closest geographical site, as necessary, to ensure model convergence.

4.2.4. Adjustments for Covariates

The analyses may be adjusted for the randomization stratification factor, prior treatment with TNF- α antagonist. It will be used in the stratified primary efficacy analysis, and as a covariate in models used for continuous endpoints.

4.2.5. Multiple Comparisons/Multiplicity

Multiplicity will be addressed via a closed testing procedure with combined p-values computed for each of the two stages of the trial (interim analysis and post-interim analysis). Each stage p-value will be adjusted as appropriate by a closed testing multiple comparison approach that controls the overall type 1 error at alpha equal to 0.025, one-sided. Further details will be provided in Section 4.6.

Testing of secondary endpoints will only be performed if the primary efficacy null hypothesis is rejected. A hierarchical testing procedure will be used for the multiple secondary endpoints at the 0.025 (one-sided) alpha level. The first ten secondary endpoints will be tested in the order that they are presented in Section 2.5.1. For secondary endpoints compared at multiple time points, only the comparison at Week 12 will be used for hierarchical testing purposes.

4.2.6. Subpopulations

Summaries on the primary and secondary efficacy endpoints will be presented by the randomization stratum, prior treatment with TNF-α antagonist.

4.2.7. Withdrawals, Dropouts, Loss to Follow-up

For subjects who are withdrawn by the Investigator (or designee) or who voluntarily withdraw prematurely from the study, replacement subjects will be enrolled only if deemed necessary by the sponsor. Replacement subjects will be assigned a subject number by adding 300 to the number of the subject they are replacing (e.g., Subject No. 305 replaces Subject No. 005) and will be enrolled into the same randomization arm as the subject whom they are replacing.

Analyses will utilize all available data obtained from such subjects (withdrawn and replacement subjects), in accordance with the study population definitions.

4.2.8. Missing, Unused, and Spurious Data

All data recorded on the CRF will be included in data listings that will accompany the clinical study report.

For the primary efficacy endpoint and other success/failure secondary endpoints, subjects who discontinue prior to Week 12 will be considered treatment failures, a worst-case scenario.

These endpoints include

- Clinical Remission (primary endpoint)
- Clinical Response
- Endoscopic Response
- Endoscopic Remission

A single sensitivity analysis on each of these endpoints will utilize the per protocol population (complete-case) in order to examine the robustness of the results.

For all continuous secondary endpoints, data will be analyzed using a mixed model repeated measures approach described in Section 4.6.2. The MMRM method assume that missing data arose from a missing at random (MAR) mechanism, i.e. the missing data is independent of unobserved data given the non-missing data. The MAR assumption is reasonable in clinical trials as the observed data explain much of the missingness in many scenarios, particularly in well controlled studies such as clinical trials, in which an extensive effort is made to observe all the outcomes and the factors that influence them. In MMRM, the information from the observed data is used via the within-patient correlation structure to provide information about unobserved data. The data from all visits are simultaneously analyzed using restricted maximum likelihood methods and no explicit imputation for missing values is performed. This method will be used in the testing of the secondary endpoints.

In addition, sensitivity analyses may be performed on complete datasets where missing data is imputed using the following methodology:

- LOCF (last observation carried forward)
- WOCF (missing data imputed using the worst observation observed prior to that time point)
- Per Protocol population (complete-case scenario- no imputation required)
- Multiple Imputation

For multiple imputation, two approaches may be used. In the first approach, for missing response data at various time points, the following steps are taken:

- Draw a random sample from the normal distribution with a mean and variance calculated from the non-missing responses at that time point
- Repeat this process five times to form five imputed complete datasets
- Fit the MMRM model to each imputed dataset to estimate the treatment effect on the response variable and the variance

- Combine the results from the five imputed dataset using SAS Proc MIANALYZE to derive the multiple imputation estimates of treatment effect and variance.

The second multiple imputation approach will be used to impute data in the event that missing data are missing not at random (MNAR) but rather depend on the observed and unobserved data. This imputation approach would use the same steps as above, but the random sample would be drawn from a distribution pre-defined and related to the responses that are observed.

Exploratory endpoints will not be imputed.

Adverse event (AE) dates that are missing or incomplete will be handled in the following manner to determine whether an AE is treatment-emergent:
If the start date/time of an AE is partially or completely missing, the date/time will be compared as far as possible with the date/time of the start of administration of study drug. The AE will be assumed to be treatment emergent if it cannot be definitively shown that the AE did not occur or worsen during the treatment-emergent period (worst case approach).

The following general rules will be used:

- If the start time of an AE is missing but the start date is complete, an AE will only be excluded as being treatment emergent if the start date is before the date of study drug administration or if the stop date/time is before study drug administration.
- If the start time and day are missing but the start month and year are complete, an AE will only be excluded as being treatment emergent if the start month/year is before the month/year of study drug administration or if the stop date/time is before study drug administration.
- If the start day and month are missing but the start year is complete, an AE will only be excluded as being treatment emergent if start year is before the year of study drug administration or if the stop date/time is before study drug administration.
- If the start date is completely missing, an AE will be considered treatment-emergent unless the stop date/time is before study drug administration.

4.2.9. Visit Windows

Table 3 below presents the allowable visit windows and intervals for analysis of efficacy variables. For the purposes of statistical analysis, the visit days will be recalculated in terms of study days since the first day of dosing as illustrated below.

Baseline will be defined as the last non-missing pre-treatment assessment. For post-baseline visits, if a subject has more than one assessment occurring in the same visit window, the data from the visit closest to the scheduled study day will be used. If two visits have the same distance from the scheduled study day, the data of the visit after the scheduled study day will be used.

Table 1 Evaluation Intervals for Efficacy Analysis

Evaluation	Protocol-Specified Interval	Interval for Analysis
Baseline (Screening and Pre-Dose)	Day –35 to Day 0	Day –35 to Day 0 (Pre-Dose)
Day 0 (Post-Dose)	Day 0 (Post-Dose)	Day 0 (Post-Dose)
Day 14	Day 11 post to Day 17 post	Day 1 post to Day 24 post
Day 28	Day 25 post to Day 31 post	Day 25 post to Day 38 post
Day 42	Day 39 post to Day 45 post	Day 39 post to Day 52 post
Day 56	Day 53 post to Day 59 post	Day 53 post to Day 66 post
Day 70	Day 67 post to Day 73 post	Day 67 post to Day 80 post
Day 84	Day 81 post to Day 87 post	Day 81 post to Day 104 post
Day 112 (Follow-Up)	Day 105 post to Day 119 post	Day 105 post to Day 119 post

Actual dates and times will be used for pharmacokinetic and pharmacodynamic analyses rather than nominal days and times.

4.3. Interim Analyses

An unblinded interim analysis (IA) will be performed by the ADRC after approximately 60-80 subjects have completed the full 12 weeks dosing. A futility analysis will be performed; if the trial does not meet criteria for futility, further analysis will be conducted in order to drop one (or two) of the PTG-100 dose arms. Following the IA, randomization will continue to the remaining dose arms (Stage 2).

Conditional power will also be computed at the IA. It is the probability that the trial will yield a statistically significant difference for clinical remission at the higher chosen dose for Stage 2 if the observed difference at the IA is the TRUE underlying difference and the trial is continued to completion. If this probability is less than 10%, the ADRC could suggest that the trial be stopped for futility after review of safety and other efficacy and exploratory endpoints. Conditional power for an active dose versus placebo is derived as follows.

Let $\hat{\delta}_2$ be the observed difference of proportions between the active dose and placebo group based on data from the second stage. Then $p_2 = \Phi\left(\frac{\hat{\delta}_2}{SE(\hat{\delta}_2)}\right)$. Therefore

$$\Phi\left(\frac{\hat{\delta}_2}{SE(\hat{\delta}_2)}\right) \leq \Phi\left\{\frac{\Phi^{-1}(\alpha) - w_1\Phi^{-1}(\tilde{p}_1)}{w_2}\right\}$$

or equivalently

$$\frac{\hat{\delta}_2}{SE(\hat{\delta}_2)} \leq \frac{\Phi^{-1}(\alpha) - w_1\Phi^{-1}(\tilde{p}_1)}{w_2}$$

Denote the true difference of response rates between the selected active dose and placebo group by δ. Then we have

$$\frac{\hat{\delta}_2 - \delta}{SE(\hat{\delta}_2)} \leq \frac{\Phi^{-1}(\alpha) - w_1\Phi^{-1}(\tilde{p}_1)}{w_2} - \frac{\delta}{SE(\hat{\delta}_2)}$$

Note that $\frac{\hat{\delta}_2 - \delta}{SE(\hat{\delta}_2)}$ follows an approximately standard normal distribution. Hence the conditional power is given by

$$P\left\{\frac{\hat{\delta}_2 - \delta}{SE(\hat{\delta}_2)} \leq \frac{\Phi^{-1}(\alpha) - w_1\Phi^{-1}(\tilde{p}_1)}{w_2} - \frac{\delta}{SE(\hat{\delta}_2)}\right\}$$

$$= \Phi\left\{\frac{\Phi^{-1}(\alpha) - w_1\Phi^{-1}(\tilde{p}_1)}{w_2} - \frac{\delta}{SE(\hat{\delta}_2)}\right\}$$

where $SE(\hat{\delta}_2) = \sqrt{\frac{\pi_t(1-\pi_t)}{n_{2t}} + \frac{\pi_c(1-\pi_c)}{n_{2c}}}$ and n_{2t}, n_{2c} are the number of subjects for the active dose and placebo group at the second stage respectively. Since we don't know δ, π_t, π_c, we will replace them by those observed from the first stage data in the conditional power calculation.

PD, efficacy, and safety data will be analyzed to model dose- and exposure-response from Stage 1 of the trial to allow dropping of an ineffective or redundant arm(s). It is also possible that a new dose arm (not tested in Stage 1 of the trial) could be added in Stage 2 for exploratory purposes. Since the trial provides for a 12-week treatment period, it is expected that an additional approximately 60 subjects would be randomized in equal proportions to the 4 treatment groups while the first approximately 60-80 subjects' 12-week data are accumulated and analyzed. Hence, in Stage 2 the remaining subjects would be randomized in equal proportions to the selected dose(s) and placebo. This would yield approximately 70-80 subjects per treatment group for the selected doses. The final analysis will combine all observed data from both Stages 1 and 2.

Dose selection will further be made in consideration of safety data and sound clinical judgment, and will be made by the ADRC), which will be comprised of the DMC and a representative of the Sponsor (not involved in the direct conduct of the study). No pre-specified algorithm will be defined for the IA. Only the ADRC, which will make the dose selection for Stage 2, and the Independent Statistical Center, which will perform the IA, will be unblinded.

4.4. Subject Disposition

Subject disposition will be tabulated, including the number screened, the number randomized by treatment group, the number with major protocol violations, the number in each subject population for analysis, the number that withdrew prior to completing the study, and reasons for withdrawal.

The disposition by study visit tabulation will be based on the FAS population in order to report specifically on the disposition of dosed subjects with endpoint data.

Disposition will also be presented graphically in the study report using a CONSORT diagram, utilizing frequencies from validated tabular outputs.

A by-subject listing of study completion information, including the reason for premature study withdrawal, if applicable, will be presented.

4.5. Demographic and Baseline Characteristics

Baseline, demographic, medical history, and physical examination at screening information will be summarized for the safety population using descriptive statistics. No formal statistical comparisons will be performed. These variables will also be summarized by analysis stage (1 vs. 2), both overall and restricted to the placebo group. Baseline disease characteristics from the sigmoidoscopy, biopsy, and a baseline Complete Mayo Score will also be summarized by analysis stage, both overall and restricted to the placebo group.

Demographic and Baseline data will be provided in data listings.

A by-subject listing of all medical history conditions and physical examination documented at screening for the randomized population will be provided.

Exposure to study drug will be summarized by treatment group for the safety population. Exposure will be summarized by total amount of time in the study and total number of days on study drug. A by-subject listing of exposure to study drug for all randomized subjects will also be provided.

4.6. Efficacy Evaluations

4.6.1. Primary Efficacy Analysis

The primary efficacy endpoint for this study is the proportion of subjects receiving PTG-100 with clinical remission at Week 12 compared with placebo, as defined in Section 2.5.1.

Efficacy analysis for the primary endpoint will be conducted using the FAS population and will be repeated on the PP population as a sensitivity analysis.

The proportion of subjects with clinical remission at Week 12 will be calculated and reported with a two-sided 95% exact confidence interval as an estimate of the true proportion of subjects with clinical remission in the target patient population. Two approaches will be used to report the percentages:

Approach 1: The naïve percentage and associated 95% CI for each dose group, based on data combined across both stages of the trial.

Approach 2: For doses that are selected for Stage 2, the Stage 2-only proportion of subjects with clinical remission and 95% CI will be reported. For doses not selected for Stage 2, the proportions and associated 95% CI based on Stage 1 data only will be reported.

Analysis to examine treatment effect (i.e. odds ratio for clinical remission among subjects with each active dose vs. placebo and 95% CI) will be carried out at the Week 12 time point, employing a Cochran-Mantel-Haenszel test stratified for prior TNF-α use (yes/no). Raw p-values, unadjusted for multiple comparisons, associated with each odds ratio will also be presented.

To obtain the primary endpoint study p-values, Simes-adjusted p-values[8] for each dose regimen vs. placebo that are produced for the primary endpoint from the first and second stages of the study will be combined by using the inverse normal combination function to obtain the final p-value.

The components contributing to the Stage 1 portion of the p-value will include data of subjects from all 4 original dose groups (placebo, low, medium, high dose) who were enrolled during Stage 1, defined as on or before the data snapshot date for the interim look at efficacy.

The components contributing to the Stage 2 portion of the p-value will be data from subjects from the dose groups that were not discontinued who were enrolled after the date of the data snapshot for the interim look at efficacy.

The final p-value for the trial primary endpoint will be a weighted combination of the two components, as described below.

The null hypothesis to be tested is

$$H_0: \pi_C = \pi_L = \pi_M = \pi_H$$

against the alternative hypothesis

$$H_1: \text{at least one of } \pi_L, \pi_M, \pi_H \text{ is different from } \pi_C$$

where π_C is the response rate for placebo arm, and π_L, π_M, π_H are the response rates for the three active dose groups. Let $w_1 = \sqrt{t_1}$ and $w_2 = \sqrt{1 - t_1}$ be the prespecified weights where t_1 is the statistical information time for interim analysis. Let p_{1L}, p_{1M}, p_{1H} be the marginal p-values for the three comparisons based on the test statistics of difference of proportions between active dose and placebo group. The test is left tailed. Let $p_{1(1)} \leq p_{1(2)} \leq p_{1(3)}$ be the ordered p-values. Let $\tilde{p}_1$ be the Simes adjusted p-value based on the first stage data. Then

$$\tilde{p}_1 = \min\{3p_{1(1)}, \frac{3}{2}p_{1(2)}, p_{1(3)}\}.$$

Let p_2 be the p-value for the comparison for the selected dose versus placebo group. The final combined Simes p-value is given by

$$\Phi(w_1\Phi^{-1}(\tilde{p}_1) + w_2\Phi^{-1}(p_2))$$

where $\Phi(.)$ is the cumulative distribution function of standard normal variable and $\Phi^{-1}(.)$ is the inverse cumulative distribution function of standard normal variable.

The final analysis is significant if

$$\Phi\big(w_1\Phi^{-1}(\tilde{p}_1) + w_2\Phi^{-1}(p_2)\big) \leq \alpha$$

which is the same as

$$p_2 \leq \Phi\left\{\frac{\Phi^{-1}(\alpha)-w_1\Phi^{-1}(\tilde{p}_1)}{w_2}\right\}$$

Analyses of the primary endpoint will be conducted separately for the subjects with prior TNF-alpha use and those without prior TNF-alpha use. Treatment effect estimates will be reported using the naïve odds ratio estimator and its associated 95% CI. P-values will be derived from unstratified Cochran-Mantel-Haenszel tests.

4.6.2. Secondary Efficacy Analyses

Analyses for the secondary endpoints will be conducted using the FAS and PP populations. Where applicable, baseline is defined as the last non-missing pre-treatment assessment prior to initiation of study treatment.

All secondary endpoints that reflect a dichotomous response will be analyzed in the same manner as the primary efficacy endpoint. These secondary endpoints include

- Proportion of subjects with endoscopic response at Week 12 (Day 84) (defined as an endoscopic subscore of 0 or 1)
- Proportion of subjects with clinical response at Week 12 (Day 84) (defined as at least 1 point and 30% reduction from baseline in rectal bleeding and stool frequency subscores)
- Proportion of subjects developing neutralizing antidrug antibodies (ADA)
    - Methods for detection of ADA in clinical samples are presented in Section 7.8 of the protocol
- Proportion of subjects with endoscopic remission at Week 12 (defined as an endoscopic subscore of 0)

Hypothesis testing for all continuous secondary endpoints will be conducted using repeated measures mixed analysis of covariance model adjusted for treatment, time point, treatment-by-time point interaction, stratification factor (TNF-alpha prior use) and baseline values. For subjects who withdraw prematurely, the last observation will be carried forward. These analyses will be based on the Stage 1 and Stage 2 combined data for exploratory purposes. Pairwise comparisons of active drug versus placebo for the change from baseline endpoints will be conducted and p-values derived as described in the analysis of the primary efficacy endpoint. In addition, a summary of observed values and change from baseline values will be presented for each endpoint. Graphical summaries by treatment group of mean change from baseline over time up to Week 12 will also be presented.

Continuous secondary endpoints include

- Mean change in complete Mayo Score (including all four subscores) from baseline to Week 12
- Mean change in Partial Mayo Score (excluding endoscopy subscore) from baseline to Weeks 2, 4, 6, 8, 10, 12, and 16 (Days 14, 28, 42, 56, 70, 84, and 112)
- Mean change in endoscopy subscore from baseline to Week 12 (Day 84)
- Mean change in rectal bleeding and stool frequency subscores from baseline to Weeks 2, 4, 6, 8, 10, 12, and 16 (Days 14, 28, 42, 56, 70, 84, and 112)
- Mean change in fecal calprotectin levels from baseline to Weeks 6 and 12
- Mean change in IBDQ score from baseline to week 12

4.6.3. Exploratory Analyses

Analyses for the exploratory endpoints will be conducted using the FAS and PP populations.

Exploratory endpoints include:

- Mean change in histological score from baseline to Week 12 (Day 84)
- Effects of ADA on PK, safety, and efficacy in subjects with positive ADA

For the first two exploratory endpoints, these will be analyzed in the same fashion as their primary and secondary counterparts. The effects of ADA on PK, safety, and efficacy may be evaluated by performing subgroup analyses (ADA + vs. ADA -) if ADA is detected in clinical samples.

4.7. Pharmacokinetic and Pharmacodynamic Evaluations

Pharmacokinetic analyses will be conducted using the Pharmacokinetic Population. Pharmacodynamic analyses will be conducted using the FAS and PP populations.

4.7.1. Pharmacokinetic Evaluations

Pharmacokinetic variables will be calculated from the individual PTG-100 plasma concentration data using standard non-compartmental methods. Maximum concentration (C_{max}), trough (predose) concentration (C_{trough}) and the time of C_{max} (T_{max}) will be taken directly from the observed data. The area under the concentration-time curve (AUC_{0-t}) from time 0 to the time of the last sample collection will be calculated using the linear trapezoidal rule.

PK analyses will be performed using commercial software such as Phoenix™ WinNonlin® Version 6.4 or higher (Certara USA Inc.). All pharmacokinetic calculations will be performed using published non-compartmental relationships. Actual dose administration times and sample collection times will be used for the analyses as recorded on the CRF. Plasma concentrations below the lower limit of quantification will be set to zero for the analysis.

Individual PTG-100 concentration data will be listed and summarized by treatment group with descriptive statistics (sample size [N], arithmetic mean, standard deviation [SD], median, minimum, maximum, geometric mean, and geometric coefficient of variation [CV]). Mean and individual PTG-100 concentration-time profiles will be presented graphically on both linear and semi-logarithmic scales. All statistical analyses will use non-rounded parameter estimates. Plasma concentration data and the pharmacokinetic parameter estimates will be presented by study group. Pharmacokinetic parameter estimates will be three or four significant figures for presentation. No attempt will made to estimate missing data. Other parameters and data handling procedures may be added as appropriate. Missing values will not be imputed, and if sufficient data are missing for a given subject, that subject may be considered non-evaluable for pharmacokinetic analysis and would not be included in the PK Population.

All plasma concentration data and the per-subject pharmacokinetic parameters will be displayed in listings.

4.7.2. Pharmacodynamic Evaluations

Quantitation of α4β7 %RO and α4β7 RE on peripheral blood T cells will be evaluated as secondary endpoints.

Exploratory endpoints will include:

1. Analysis of peripheral blood lymphocytes in subjects on individual PTG-100 dose levels to placebo
2. Average number of β7 positive cells in colonic biopsies as assessed by immunohistochemistry (IHC) in subjects by PTG-100 dose levels and compared to placebo
3. Characterization of immunologic biomarkers in the target population (selected sites only)

Individual data will be listed for each individual and summarized by nominal sampling time point and treatment group with descriptive statistics (arithmetic mean, SD, median, minimum, and maximum). A summary of change-from-baseline at each protocol specified time-point by treatment group will also be presented. The change of PD indicators over time will be demonstrated graphically and will be compared among the dose levels using the same mixed model as the other continuous secondary and exploratory endpoints.

Further summaries may use pre-clinical and clinical data from previous studies to explore the PK/PD relationship to correlate with clinical endpoints.

Receptor occupancy and α4β7 expression in peripheral blood will be listed for each individual and summarized by treatment group using frequency/count data.

4.8. Safety Analyses

Safety analyses will be conducted using the Safety Population. The safety analyses generated will be a presentation of AEs, SAEs, clinical labs, vital signs, ECGs, and physical examinations as described below.

4.8.1. Adverse Events

The primary safety endpoint is the proportion of subjects with at least 1 AE comparing individual PTG-100 dose groups with placebo.

Secondary endpoints (all based on comparison of individual PTG-100 dose levels to placebo) will include:

- Frequency and type of AEs (affecting ≥ 5% of subjects)
- Proportion of subjects with at least 1 SAE
- Number and type of SAEs
- Frequency of adverse events of special interest (AESI) including serious or opportunistic infection (viral, bacterial, fungal including systemic/ gut localization), allergic/ drug reactions, immune system disorders, and PML or new onset neurologic symptoms

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and displayed in tables and listings using System Organ Class (SOC) and Preferred Term.

Analyses of adverse events will be performed for those events that are considered treatment emergent, where treatment emergent is defined as any adverse event with onset after the administration of study medication through the end of the study visit or any event that was present at baseline but worsened in intensity or was subsequently considered drug-related by the investigator through the end of the study.

Adverse events are summarized by subject incidence rates, therefore, in any tabulation, a subject contributes only once to the count for a given adverse event (SOC or preferred term).

The number and percentage of subjects

- with any treatment-emergent adverse event
- with any treatment-emergent adverse events assessed by the Investigator as related to treatment (definite, probable, or possible relationship)
- with any treatment-emergent serious adverse event
- with treatment-emergent adverse events by severity
- with adverse events affecting ≥ 5% of subjects within any treatment group across both stages of the study
- with AESI

will be summarized by treatment group and overall. In addition, all active groups will be pooled in the summaries for comparison to placebo. In these tabulations, each subject will contribute

only once (i.e., the most related occurrence or the most intense occurrence) to each of the incidence rates in the descriptive analysis, regardless of the number of episodes.

Three additional tables will present the incidence of treatment-emergent adverse events by PT, treatment-emergent events leading to study withdrawal and treatment-emergent events with a fatal outcome, sorted in the order of frequency for the entire safety population.

All adverse events occurring on study (pre-treatment, treatment, follow-up) will be listed in subject data listings.

By-subject listings also will be provided for the following: subject deaths; serious adverse events; and adverse events leading to withdrawal.

4.8.2. Laboratory Data

Clinical laboratory values will be expressed using SI units.

A tabulation (n, %) by treatment group and overall of whether a subject had a clinical abnormality at any time, and a listing of clinical abnormalities, for all identified laboratory abnormalities, will be generated.

Tabulation of actual value and change from baseline to each on study evaluation will be presented for each clinical laboratory parameter, including hematology and clinical chemistry. In the event of repeat values, the last non-missing value per study day/time will be used.

Continuous lab parameters will be categorized as low, medium, or high based on the value of the parameter. Shift tables for lab parameters from baseline to all applicable visits will be presented.

All laboratory data will be provided in data listings.

4.8.3. Vital Signs and Physical Examinations

Tabulation of actual value and change from baseline to each on study evaluation will be presented for all vital signs. No hypothesis testing will be conducted for vital sign data.

By-subject listings of vital sign measurements will be presented in data listings.

Physical examination results at each time point will be summarized; shifts from baseline in physical examination findings to each on study visit will also be presented. Clinically significant changes in physical examinations will be listed for each participant and described in the text of the final report.

A tabulation (n, %) of subjects with a physical examination abnormality will be presented by visit and dose group. The tabulation will be presented by body system. Subjects with no examination for a particular body system will not be included in the summary statistics for that body system. No formal statistical comparisons will be performed.

All physical examination findings will be presented in a data listing.

4.8.4. Electrocardiogram

Tabulation of actual value and change from baseline to each on study evaluation will be presented for all continuous ECG measurements (heart rate, PR interval, QRS interval, QT interval). No hypothesis testing will be conducted for vital sign data.

ECG results will also be summarized by the number and percent of subjects with overall normal, abnormal and clinically significant abnormal results at baseline and each study visit.

All ECG data for each subject will be provided in data listings.

4.8.5. Concomitant Medications

Concomitant medications will be coded using the WHO Drug dictionary. Results will be tabulated by Anatomic Therapeutic Class (ATC) and preferred term.

The use of concomitant medications will be included in by-subject data listing.

5. CHANGES TO PLANNED ANALYSES

As of this date, there have been no changes between the protocol-defined statistical analyses and those presented in this statistical plan.

6. REFERENCES

1. Rutgeerts P, Sandborn WJ, Feagan BG, Reinisch W, Olson A, Johanns J et al. (2005). Infliximab for induction and maintenance therapy for ulcerative colitis. N Engl J Med 353(23):2462-2476 & [Correction, 2006, 354(20):2200].

2. Sandborn WJ, Van Assche G, Reinisch W, Colombel J-F, D'Haens G, Wolf DC et al. (2012). Adalimumab induces and maintains clinical remission in patients with moderate-to-severe ulcerative colitis. Gastroenterology 142(2):257-265.e3.

3. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel J-F, Sandborn WJ et al. (2013). Vedolizumab as induction and maintenance therapy for ulcerative colitis GEMINI 1 Study Group. N Engl J Med 369(8):699-710.

4. Sandborn WJ, Feagan BG, Rutgeerts P, Hanauer S, Colombel J-F, Sands BE et al. (2013). Vedolizumab as induction and maintenance therapy for Crohn's disease GEMINI 2 Study Group. N Engl J Med 369(8):711-721.

5. Takeda Pharmaceuticals America, Inc. (2014). Prescribing Information: ENTYVIO (Vedolizumab) for Injection, for Intravenous Use. [Label Package]. Deerfield (IL): Takeda Pharmaceuticals America, Inc. Available at: http://general.takedapharm.com/content/file.aspx?FileTypeCode=ENTYVIOPI&cacheRandomizer=1b94c181-3ea4-4aa4-8ac5-3a0592873792 [May, 2014].

6. Gibaldi M, Perrier D. Pharmacokinetics. 2nd edition. New York, NY: Marcel Dekker Inc.; 1982.

7. Department of Health and Human Services, Food and Drug Administration. International Conference on Harmonisation; Guidelines on Structure and Content of Clinical Study Reports; Availability. Federal Register. 1996;61(138):37320-37343.

8. Simes RJ. An improved Bonferroni procedure for multiple tests of significance. Biometrika. 1986;73:751-754

7. CLINICAL STUDY REPORT APPENDICES

7.1. Statistical Tables to be Generated

Disposition/Demographics/Baseline/Exposure (CSR Table Section 14.1)

Table 14.1.1.1	Subject Enrollment and Disposition (All Subjects)
Table 14.1.1.2	Subject Disposition by Visit (FAS)
Table 14.1.2	Demographic and Baseline Characteristics (Safety Population)
Table 14.1.3	Baseline Disease Characteristics (Safety Population)
Table 14.1.4	Medical History (Safety Population)
Table 14.1.5	Physical Examination at Screening (Safety Population)
Table 14.1.6	Exposure to Study Treatment (Safety Population)

Efficacy/Pharmacokinetic Results (CSR Table Section 14.2)

Table 14.2.1.1A	Proportion of Subjects with Clinical Remission at Week 12 (FAS)
Table 14.2.1.1B	Proportion of Subjects with Clinical Remission at Week 12 (PP)
Table 14.2.1.2A	Proportion of Subjects with Clinical Remission at Week 12 by Prior TNF-alpha Use (FAS)
Table 14.2.1.2B	Proportion of Subjects with Clinical Remission at Week 12 by Prior TNF-Alpha Use (PP)
Table 14.2.1.3A	Proportion of Subjects with Endoscopic Response at Week 12 (FAS)
Table 14.2.1.3B	Proportion of Subjects with Endoscopic Response at Week 12 (PP)
Table 14.2.1.4A	Proportion of Subjects with Clinical Response at Week 12 (FAS)
Table 14.2.1.4B	Proportion of Subjects with Clinical Response at Week 12 (PP)
Table 14.2.1.5A	Proportion of Subjects with Endoscopic Remission at Week 12 (FAS)
Table 14.2.1.5B	Proportion of Subjects with Endoscopic Remission at Week 12 (PP)
Table 14.2.2.1A	Change in complete Mayo Score from Baseline to Week 12 (FAS)

Table 14.2.2.1B Change in complete Mayo Score from Baseline to Week 12 (PP)

Table 14.2.2.1C Change in complete Mayo Score from Baseline to Week 12 (Sensitivity Analyses - if Needed)

Table 14.2.2.2A Change in Partial Mayo Score from Baseline Over Time (FAS)

Table 14.2.2.2B Change in Partial Mayo Score from Baseline Over Time (PP)

Table 14.2.2.2C Change in Partial Mayo Score from Baseline Over Time (Sensitivity Analyses - if Needed)

Table 14.2.2.3A Change in Endoscopy Subscore from Baseline to Week 12 (FAS)

Table 14.2.2.3B Change in Endoscopy Subscore from Baseline to Week 12 (PP)

Table 14.2.2.3C Change in Endoscopy Subscore from Baseline to Week 12 (Sensitivity Analyses – if Needed)

Table 14.2.2.4A Change in Rectal Bleeding Subscore from Baseline Over Time (FAS)

Table 14.2.2.4B Change in Rectal Bleeding Subscore from Baseline Over Time (PP)

Table 14.2.2.4C Change in Stool Frequency Subscore from Baseline Over Time (Sensitivity Analyses – if Needed)

Table 14.2.2.5A Change in Stool Frequency Subscore from Baseline Over Time (FAS)

Table 14.2.2.5B Change in Stool Frequency Subscore from Baseline Over Time (PP)

Table 14.2.2.5C Change in Stool Frequency Subscore from Baseline Over Time (Sensitivity Analyses – if Needed)

Table 14.2.2.6A Change in Fecal Calprotectin Levels from Baseline Over Time (FAS)

Table 14.2.2.6B Change in Fecal Calprotectin Levels from Baseline Over Time (PP)

Table 14.2.2.6C Change in Fecal Calprotectin Levels from Baseline Over Time (Sensitivity Analyses – if Needed)

Table 14.2.2.7A Change in IBDQ score from Baseline Over Time (FAS)

Table 14.2.2.7B Change in IBDQ score from Baseline Over Time (PP)

Table 14.2.2.7C	Change in IBDQ score from Baseline Over Time (Sensitivity Analyses – if Needed)
Table 14.2.2.7A	Change in Histological score from Baseline Over Time (FAS)
Table 14.2.2.7B	Change in Histological score from Baseline Over Time (PP)
Table 14.2.3	Summary of Pharmacokinetic Parameters (PK Population)
Table 14.2.4.1A	Change in α4β7 Receptor Occupancy on Peripheral Blood T cells Over Time (FAS)
Table 14.2.4.1B	Change in α4β7 Receptor Occupancy on Peripheral Blood T cells Over Time (PP)
Table 14.2.4.2A	Summary of α4β7 Receptor Expression on Peripheral Blood T cells Over Time (FAS)
Table 14.2.4.2B	Summary of α4β7 Receptor Expression on Peripheral Blood T cells Over Time (PP)
Table 14.2.4.3A	Change in Peripheral Blood Lymphocytes Over Time (FAS)
Table 14.2.4.3B	Change in Peripheral Blood Lymphocytes Over Time (PP)
Table 14.2.4.4A	Change in β7 Positive Cells in Colonic Biopsies Over Time (FAS)
Table 14.2.4.4B	Change in β7 Positive Cells in Colonic Biopsies Over Time (PP)
Table 14.2.5.1A	Proportion of Subjects with Positive Neutralizing ADA by Week 12 and Week 16 (FAS)
Table 14.2.5.1B	Proportion of Subjects with Positive Neutralizing ADA by Week 12 and Week 16 (PP)

Displays of Adverse Events (CSR Section 14.3.1)

Table 14.3.1.1	Overall Summary of Treatment-Emergent Adverse Events by Treatment Group (Safety)
Table 14.3.1.2	Summary of Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Treatment Group, Sorted in Descending Frequency according to the Total Column (Safety)

Table 14.3.1.3 Summary of Related/Possibly, Probably and Definitely Related Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Treatment Group (Safety)

Table 14.3.1.4 Summary of Treatment-Emergent Adverse Events by Severity, System Organ Class, Preferred Term and Treatment Group (Safety)

Table 14.3.1.5 Summary of Treatment-Emergent Serious Adverse Events by System Organ Class, Preferred Term and Treatment Group (Safety)

Table 14.3.1.6 Summary of Most Common (>=5% within any treatment group) Treatment-Emergent Adverse Events by System Organ Class, Preferred Term and Treatment Group (Safety)

Table 14.3.1.7 Summary of Treatment-Emergent Adverse Events by Preferred Term and Treatment Group (Safety)

Table 14.3.1.8 Summary of Adverse Events Leading to Study Withdrawal by System Organ Class, Preferred Term and Treatment Group (Safety)

Table 14.3.1.9 Summary of Treatment-Emergent Adverse Events with a Fatal Outcome by System Organ Class, Preferred Term and Treatment Group (Safety)

Listings of Deaths, Serious Adverse Events and Other Significant Events (CSR Section 14.3.2)

Table 14.3.2.1 Listing of Subject Deaths

Table 14.3.2.2 Listing of Serious Adverse Events

Table 14.3.2.3 Listing of Adverse Events Leading to Withdrawal or Death

Abnormal Laboratory Value Tables and Listings (CSR Section 14.3.4)

Table 14.3.4.1.1 Summary of Clinical Laboratory Abnormalities: Hematology (Safety)

Table 14.3.4.1.2 Summary of Clinical Laboratory Abnormalities: Chemistry (Safety)

Table 14.3.4.1.3 Summary of Clinical Laboratory Abnormalities: Urinalysis (Safety)

Table 14.3.4.1.4 Summary of Clinical Laboratory Abnormalities: Coagulation (Safety)

Table 14.3.4.2.1 Listing of Clinical Laboratory Abnormalities: Hematology (Safety)

Table 14.3.4.2.2 Listing of Clinical Laboratory Abnormalities: Chemistry (Safety)

Table 14.3.4.2.3 Listing of Clinical Laboratory Abnormalities: Urinalysis (Safety)

Table 14.3.4.2.4 Listing of Clinical Laboratory Abnormalities: Coagulation (Safety)

Other Safety Information (CSR Section 14.4)

Table 14.4.1.1 Clinical Laboratories: Hematology Actual Values and Changes Over Time on Study (Safety Population)

Table 14.4.1.2 Clinical Laboratories: Chemistry Actual Values and Changes Over Time on Study (Safety Population)

Table 14.4.1.3 Clinical Laboratories: Urinalysis Actual Values and Changes Over Time on Study (Safety Population)

Table 14.4.1.4 Clinical Laboratories: Coagulation Actual Values and Changes Over Time on Study (Safety Population)

Table 14.4.2.1 Clinical Laboratories: Shifts from Baseline Over Time on Study for Hematology (Safety Population)

Table 14.4.2.2 Clinical Laboratories: Shifts from Baseline Over Time on Study for Clinical Chemistry (Safety Population)

Table 14.4.2.1 Clinical Laboratories: Shifts from Baseline Over Time on Study for Urinalysis (Safety Population)

Table 14.4.2.2 Clinical Laboratories: Shifts from Baseline Over Time on Study for Coagulation (Safety Population)

Table 14.4.3 Vital Signs: Actual Values and Changes Over Time on Study (Safety Population)

Table 14.4.4 Shifts from Baseline Over Time on Study for Physical Examination Findings (Safety Population)

Table 14.4.5 Electrocardiogram Results: Actual Values and Changes Over Time on Study (Safety Population)

Table 14.4.6 Concomitant Medications (Safety Population)

Table 14.4.7 Summary of Physical Examinations

7.2. Data Listings to be Generated

Discontinued Subjects (CSR Appendix 16.2.1)

Listing 16.2.1 Subject Disposition and Study Termination Information

Protocol Deviations (CSR Appendix 16.2.2)

Listing 16.2.2.1 Inclusion Criteria

Listing 16.2.2.2 Exclusion Criteria

Listing 16.2.2.3 Admission and Discharge Dates

Listing 16.2.2.4 Protocol Deviations

Subjects Excluded from Efficacy Analysis (CSR Appendix 16.2.3)

Listing 16.2.3.1 Subjects Excluded from Efficacy Analyses

Demographics and Baseline Data (CSR Appendix 16.2.4)

Listing 16.2.4.1 Demographic and Baseline Information

Listing 16.2.4.2 Baseline Disease Data

Listing 16.2.4.3 Medical History

Listing 16.2.4.4 Hepatitis and TB Testing

Compliance and/or Drug Concentration Data (CSR Appendix 16.2.5)

Listing 16.2.5.1 Dosing Information

Listing 16.2.5.2 Pharmacokinetic and Pharmacodynamic Assay Samples

Listing 16.2.5.3 Pharmacokinetic Data

Individual Efficacy Response Data (CSR Appendix 16.2.6)

Listing 16.2.6.1 Efficacy Results: Complete Remission, Endoscopic Response and Remission, and Clinical Response

Listing 16.2.6.2	Efficacy Results: Complete and Partial Mayo Scores (Including Subscores) Over Time
Listing 16.2.6.3	Efficacy Results: Fecal Calprotectin Levels Over Time
Listing 16.2.6.4	Efficacy Results: IBDQ Scores Over Time
Listing 16.2.6.5	Efficacy Results: Histological Scores Over Time

Adverse Event Listings (each subject) (CSR Appendix 16.2.7)

Listing 16.2.7.1	Adverse Events by MedDRA SOC, Verbatim Term, Preferred Term and Subject

Listing of Individual Laboratory Measurements by Subject (CSR Appendix 16.2.8)

Listing 16.2.8.1	Laboratory Results: Hematology
Listing 16.2.8.2	Laboratory Results: Chemistry
Listing 16.2.8.3	Laboratory Results: Urinalysis
Listing 16.2.8.4	Laboratory Results: Coagulation

Other Safety Data (CSR Appendix 16.2.9)

Listing 16.2.9.1	Vital Signs
Listing 16.2.9.2	Physical Examinations
Listing 16.2.9.3	Electrocardiograms
Listing 16.2.9.4	Prior and Concomitant Medications
Listing 16.2.10	Pharmacodynamic Parameters
Listing 16.2.11	Immunogenicity (ADA Status)
Listing 16.2.12	Serum Pregnancy Testing
Listing 16.2.13	Sigmoidoscopy and Colonic Biopsy