STATISTICAL ANALYSIS PLAN

Protocol Number:	URO-901-2001
Protocol Title:	A Phase 2, Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of Vibegron Administered Orally for 12 Weeks to Women with Irritable Bowel Syndrome
Protocol Version and Date:	Version 4.0 (Amendment 3); 19 Nov 2019
Investigational Product:	Vibegron (URO-901, RVT-901, MK-4618, KRP-114V)
Indication	Irritable Bowel Syndrome (IBS)
Development Phase:	Phase 2
US IND Number:	140291
EudraCT Number:	Not Applicable
Sponsor:	Urovant Sciences, GmbH Viadukststrasse 8 4051 Basel Switzerland Telephone +41 (42) 2155999
SAP Version **Effective Date**	v2.0 21Sep2020
SAP Author	[redacted] [redacted], Biostatistics, Programming and Data Management, Urovant Sciences, Inc.

NCT Number: NCT02929329
This NCT number has been applied to the document for purposes of posting on Clinicaltrials.gov

SIGNATURE PAGE

This document has been prepared and/or reviewed by:

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I am the author of this document
Signing Time: 22-Sep-2020 | 03:24 PDT
59281D577C78441691B2543FE0854F54

22-Sep-2020 | 03:24 PDT

[redacted] Biostatistics, Programming and Data Management, Urovant Sciences, Inc

Date (dd-MMM-yyyy)

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 21-Sep-2020 | 13:27 PDT
10030D83B0DB457682147A274A8713E0

21-Sep-2020 | 13:27 PDT

[redacted] Clinical Development, Urovant Sciences, Inc

Date (dd-MMM-yyyy)

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 21-Sep-2020 | 13:15 PDT
F6C869F8DB8347A8A97B7EFE293AE332

21-Sep-2020 | 13:15 PDT

[redacted] Clinical Development, Urovant Sciences, Inc

Date (dd-MMM-yyyy)

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 21-Sep-2020 | 13:32 PDT
27011A7A23334261999CAD0B9B256828

21-Sep-2020 | 13:33 PDT

[redacted] Clinical Operations, Urovant Sciences, Inc

Date (dd-MMM-yyyy)

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 21-Sep-2020 | 14:11 PDT
[redacted]76ACC4DF79AAAB2

21-Sep-2020 | 14:12 PDT

[redacted] Urovant Safety Consultant, [redacted]

Date (dd-MMM-yyyy)

This document has been approved by

DocuSigned by:
[redacted]
Signer Name: [redacted]
Signing Reason: I approve this document
Signing Time: 22-Sep-2020 | 03:24 PDT
59281D577C78441691B2543FE0854F54

22-Sep-2020 | 03:24 PDT

[redacted] Biostatistics, Programming and Data Management
Urovant Sciences, Inc.

Date (dd-MMM-yyyy)

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Term
ADaM	analysis data model
AE	adverse event
AESI	adverse events of special interest
API	Abdominal Pain Intensity
BMI	body mass index
BP	blood pressure
CFB	change from baseline
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRF	case report form
DBP	diastolic blood pressure
DSMB	data safety monitoring board
ECG	electrocardiogram
eCRF	electronic case report form
FAS	full analysis set
GIS	global improvement scale
IBD	inflammatory bowel disease
IBS	irritable bowel syndrome
IBS-D	IBS that is associated with predominantly diarrhea
IBS-M	IBS that has mixed episodes of diarrhea and constipation
[redacted]	[redacted]
IxRS	interactive voice or web response system
LS	least squares
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	mixed model for repeated measures
NRS	Numeric Rating Scale
PT	preferred term
PRO	patient reported outcome
REML	restricted (or residual) maximum likelihood
SAE	serious adverse event
SAF	safety set
SAP	statistical analysis plan
SBP	systolic blood pressure
SD	standard deviation
SDTM	study data tabulation model
SE	standard error
SOC	system organ class
TEAE	treatment emergent adverse event

Abbreviation	Term
WAP	worst abdominal pain
WHO	World Health Organization
[redacted]	[redacted]
β3-AR	beta-3 adrenergic receptor

SAP VERSION HISTORY

Version	*Date*	*Description of Changes*
1.0	**25June 2020**	***Original Document***
2.0		***See Appendix 11.3 for a description of changes***

1. INTRODUCTION

This Statistical Analysis Plan (SAP) provides a description of the statistical methodology to be implemented for the analyses of data from Protocol Version 1.0, and to be included in the clinical study report. Any deviations from this analysis plan will be documented in the final clinical study report.

Irritable bowel syndrome is a chronic, functional gastrointestinal disorder associated with abdominal pain and changes in the frequency or consistency of stool (i.e., diarrhea, constipation, or a combination of both). The Rome IV criteria classify IBS into 4 distinct subtypes based on predominant stool consistency: IBS with predominantly constipation (IBS-C), IBS with predominantly diarrhea (IBS-D), IBS with mixed episodes of diarrhea and constipation (IBS-M), and IBS with undefined subtype (IBS-U). The definitions of diarrhea and constipation correlate with the Bristol Stool Form Scale, with Types 1 and 2 defined as constipation and Types 6 and 7 defined as diarrhea. The Rome IV IBS classification is based on the patient's predominant bowel habit on days with abnormal bowel movement as follows:

- IBS-C: ≥ 25% constipation (Bristol Type 1 or 2) and < 25% diarrhea (Bristol Type 6or 7).
- IBS-D: ≥ 25% diarrhea (Bristol Type 6 or 7) and < 25% constipation (Bristol Type 1 or 2)
- IBS-M: ≥ 25% constipation (Bristol Type 1 or 2) and ≥ 25% diarrhea (Bristol Type 6or 7)
- IBS-U: unclassified (insufficient abnormality to meet other criteria)

The Rome IV diagnostic criteria must be met within the most recent 3 months, with symptom onset at least 6 months before diagnosis. The primary population is IBS-D for this proof of concept study with a small group of subjects with IBS-M for exploratory evaluation.

1.1. Study Objectives and Endpoints

1.1.1. Primary Efficacy Objectives

Primary Efficacy Objectives	Endpoints
To compare the effect of vibegron vs placebo in subjects with abdominal pain due to IBS-D on the abdominal pain intensity (API) weekly responder rate over 12 weeks	– Proportion of IBS-D subjects who are API weekly responders with ≥30% improvement over 12 weeks • An API Weekly Responder is defined as a subject who experiences a decrease in the weekly average of "worst abdominal pain in the past 24 hours" scores of at least 30% compared with the baseline weekly average

1.1.2. Secondary Efficacy Objectives

Secondary Efficacy Objectives	Endpoints
To compare the effect of vibegron vs placebo in subjects with abdominal pain due to IBS-D or IBS-M on patient-reported outcomes	– Proportion of Global Improvement Scale (GIS) responders at Week 12 for all IBS subjects including IBS-D and IBS-M subjects.
To compare the effect of vibegron vs	– Proportion of IBS-D subjects who are API weekly

placebo in subjects with abdominal pain due to IBS-D on the abdominal pain intensity (API) weekly responder rate over 12 weeks based on different thresholds of improvement	responders with ≥40% improvement over 12 weeks • An API Weekly Responder is defined as a subject who experiences a decrease in the weekly average of "worst abdominal pain in the past 24 hours" scores of at least 40% compared with the baseline weekly average
	– Proportion of IBS-D subjects who are API weekly responders with ≥50% improvement over 12 weeks • An API Weekly Responder is defined as a subject who experiences a decrease in the weekly average of "worst abdominal pain in the past 24 hours" scores of at least 50% compared with the baseline weekly average
To compare the effect of vibegron vs placebo in subjects with abdominal pain due to IBS-D or IBS-M on safety endpoints	– AEs, clinical laboratory values, vital signs

1.1.3. Other Efficacy Objectives

Other Efficacy Objectives	Endpoints
[redacted]	

Other Efficacy Objectives	Endpoints
[redacted]	

Other Efficacy Objectives	Endpoints
[redacted]	

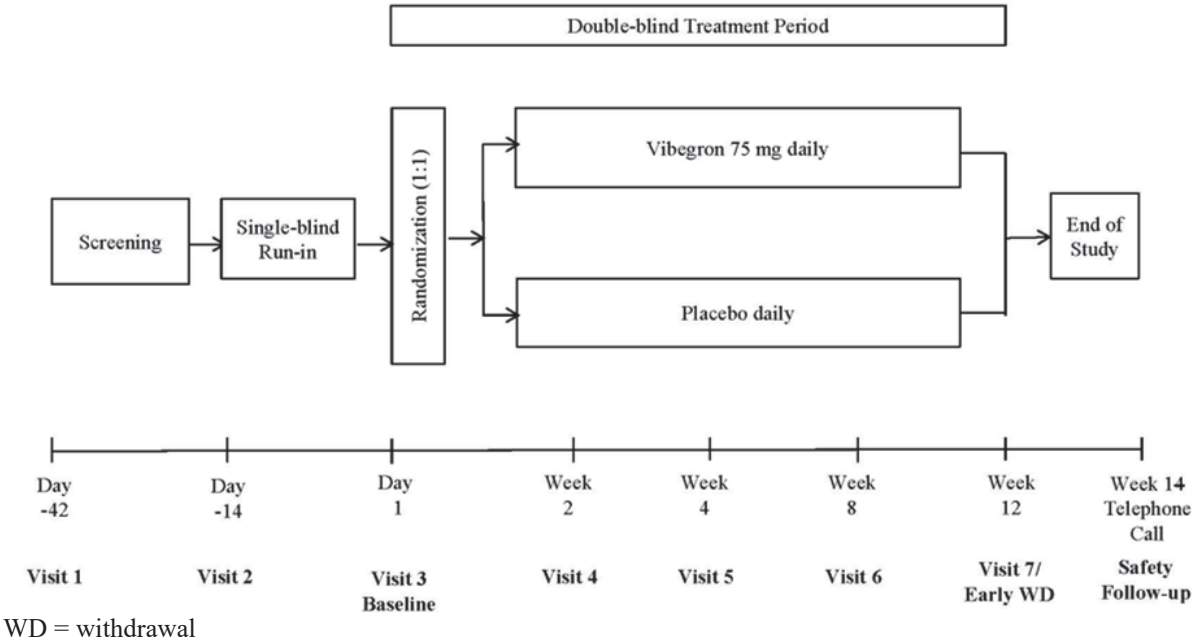
1.2. Study Design

1.2.1. Overall Study Design

This study is a Phase 2 randomized, double-blind, placebo-controlled, parallel-group, multicenter study to evaluate the efficacy and safety of vibegron in women with IBS-D or IBS-M. Subjects who meet all eligibility criteria will be randomized in a 1:1 ratio to receive either vibegron 75 mg or matched placebo. Approximately 200 subjects will be enrolled at approximately 40 sites in the US to achieve approximately 180 evaluable subjects (90 per treatment group).

The study consists of a Screening Period (1 to 5 weeks), a single-blind Placebo Run-in Period (2 weeks), a randomized, Double-blind Treatment Period (12 weeks), and a Safety Follow-up Period (2 weeks). Subjects will have a Safety Follow-up phone call approximately 14 days after the subject's last dose of study drug (i.e., at Week 14 for subjects who complete the Week 12 Visit, or approximately 2 weeks after the last dose of study drug for subjects who discontinue treatment early). The study schema is shown in Figure 1.

Figure 1: Study Schema



WD = withdrawal
The Week 14 Safety Follow-up telephone call should occur 7 to 14 days after the last dose of study drug (includes window of – 7 days) only for randomized subjects who complete Week 12 or withdraw from the study early.

Refer to Section 1.3 of the protocol for a detailed time and events schedule.

1.2.2. Randomization and Blinding

Randomization will occur centrally using an interactive voice or web response system (IxRS) using central, stratified block randomization. Randomization will be stratified based on the following stratification factors:

- Baseline abdominal pain intensity score (<6 vs ≥ 6) on a 0 to 10 numeric rating scale [NRS], and
- IBS subtype (IBS-D vs IBS-M)

For this study a double-blind/masking technique will be used: vibegron and its matching placebo will be packaged identically so that treatment blind/masking is maintained. The subject, the Investigator, and Sponsor personnel or delegate(s) who are involved in the treatment or clinical evaluation of the subjects will not be aware of the treatment group assignments.

1.2.3. Statistical Hypotheses

The primary objective will be to estimate the treatment effect of vibegron relative to placebo with respect to improvement in IBS-related abdominal pain in IBS-D subjects. There is no formal statistical hypothesis. Nominal p-values from comparisons to placebo may be provided

for descriptive purposes. An improvement in the primary endpoint is defined as a subject who experiences a decrease in the weekly average of "worst abdominal pain in the past 24 hours" scores of at least 30% compared with the baseline weekly average. Sample Size Justification

Approximately 200 subjects will be randomized in a 1:1 ratio to receive one of the following study treatments:

- Vibegron 75 mg tablet (N = 100)
- Matching placebo tablet (N = 100)

At most, 50% of randomized subjects (approximately 100 subjects total [approximately 50 per treatment arm]) will have IBS-M; at least 50% of randomized subjects (approximately 100 subjects total [approximately 50 per treatment arm]) will have IBS-D, and the IBS-D subgroup will be used for the primary endpoint analysis. Assuming a total of 10% of subjects will discontinue prior to Week 12 (for any reason), there will be a minimum of approximately 90 evaluable IBS-D subjects (approximately 45 in the vibegron arm and approximately 45 in the placebo arm) at the end of Week 12. The study has approximately 60% power to detect a between-group treatment difference of 20% in proportion of abdominal pain responders at a 2 - sided test at the $\alpha = 0.10$ level assuming a responder rate of 51% versus 31% for vibegron and placebo, respectively. The assumptions were based on results from a solabegron study female subgroup analysis [Kelleher, 2008].

2. PLANNED ANALYSES

2.1. Interim Analysis

[redacted]

2.2. Final Analysis

[redacted] will perform the production and quality control of all tables, figures and listings on behalf of Urovant Sciences, Inc.

Statistical programming will start after data have been collected and are available in the database. Blinded dry runs using dummy treatment code will be performed prior to database lock and unblinding to ensure programming displays and algorithms are developed as planned.

3. ANALYSIS POPULATION

3.1. Analysis Sets

3.1.1. Screened Set

The Screened Set consists of all subjects who are screened for the study. This population is used primarily for subject accounting purposes and will generally not be used for summary or analysis.

3.1.2. Placebo Run-in Set

The Placebo Run-in Set consists of all subjects who entered the single-blind placebo Run-in period of the study. Subjects will be considered Run-in failures if they enter the Run-in period but are not randomized to receive double-blind medication.

3.1.3. Randomized Set

The Randomized Set consists of all subjects who are randomized to receive any double-blind study medication regardless of whether they took a dose.

3.1.4. Safety Analysis Set

The Safety Analysis Set (SAF) consists of all subjects who receive at least one dose of double-blind study medication. Subjects will be classified according to the treatment they actually received.

3.1.5. Full Analysis Set

The Full Analysis Set (FAS) consists of all randomized subjects who took at least one dose of double-blind study medication and had at least one evaluable post-baseline weekly API score (i.e., where evaluable is considered minimum of 5 diary entries in a week). Subjects will be analyzed according to their randomized treatment, irrespective of whether or not they have prematurely discontinued, according to the Intent-to-Treat principle. The Full Analysis Set for IBS-D (FAS-D) and Full Analysis Set for IBS-M (FAS-M) are each a subset of FAS and consists of IBS-D and IBS-M subjects, respectively, based on the randomization strata. In general, the efficacy endpoints will be summarized using all three populations: FAS, FAS-D

and FAS-M. However, the FAS-D will serve as the primary population for the primary analysis of efficacy data in this study.

3.1.6. Per-Protocol Set

The Per-Protocol Set for IBS-D (PPS-D) excludes subjects from the FAS-D due to important deviations from the protocol that may substantially affect the result of the primary efficacy endpoint. This Per -Protocol Set will serve as the supportive population for the analysis of efficacy data in this study. Other Per Protocol Sets may be defined. The final determination on major protocol deviations, and thereby the composition of the Per-Protocol Sets, will be made prior to the unblinding of the database and will be documented.

3.2. Violations and Deviation

Subjects who do not meet eligibility criteria but were still randomized will be analyzed according to the analysis sets described in Section 3.1.

3.2.1. Protocol Deviations

The final list of major protocol deviations will be finalized and documented prior to database lock except for the deviation category of wrong treatment which will be confirmed upon study unblinding. Only major (i.e., significant per the protocol deviation plan) protocol deviations will be summarized and listed in the Clinical Study Report (CSR). Major deviations will be those which are considered to potentially impact upon the interpretation of the primary efficacy endpoint in the study or may potentially impact the interpretation of safety. Major protocol deviations may include, but are not limited to the following:

- Randomized subjects who do not meet the inclusion criteria
- Randomized subjects who meet any of the exclusion criteria
- Subjects who received the wrong treatment
- Concomitant use of prohibited medications
- Randomized subjects who met withdrawal criteria during the study but were not withdrawn
- Subjects whose IP was interrupted 14 consecutive days or more during the double-blind treatment period OR had IP compliance less than 80% from Weeks 8-12
- Subjects who were not compliant in study procedures (e.g., if 3 or more days of diary entry were missed between visits that were a 2-week period or if 6 or more days of diary entry missed between visits that were a 4-week period)

4. GENERAL STATISTICAL CONSIDERATIONS

4.1. General Principles for Data Analysis

4.1.1. Multicenter Study

In this study the stratified permuted block randomization is not done within centers. The analyses will be conducted by pooling data from all study centers and will not include study center as a covariate in the statistical modeling.

4.1.2. Testing Strategy and Multiplicity Adjustments

No formal multiplicity adjustment will be performed. All efficacy analyses will be considered descriptive. Ninety-percent confidence intervals will be provided for each treatment group and nominal p-values from comparisons to placebo may be provided for descriptive purposes as a measure of the strength of the association between the endpoint and the treatment effect.

4.1.3. Examination of Subgroups

To determine whether the treatment effect is consistent across various subgroups, the primary efficacy endpoint will be summarized descriptively for each of the following subgroups:

- Baseline abdominal pain intensity score (<6 vs ≥ 6)
- Age category (<40, ≥40 to <65, ≥65 years)
- Race (as CRF page)
- Randomized prior to 31Dec2019 vs Randomized after 31Dec2019
- Baseline Bristol Stool Scale (< 5.5 vs ≥ 5.5)

Note, the protocol amendment in Nov2019 changed the fecal calprotectin testing to optional and that it should only be considered if there is a strong suspicion in the opinion of the investigator that the subject has inflammatory bowel disease (IBD) (e.g., family history in a 1st degree relative). In order to allow time operationally for sites to implement this a date of 31Dec2019 was chosen as the cutoff.

All efficacy data including eDiary and ePRO data will be summarized by IBS sub-type (IBS-D and IBS-M) and all IBS.

4.2. General Data Handling Conventions

4.2.1. Study Treatment Description

Randomized treatment groups will be displayed as show in the following table:

Data Displays for Reporting

Description	Order in TLF
Placebo	1
Vibegron 75mg	2
Overall	3

4.2.2. Reporting Conventions

General rules

In general, all collected safety data and any derived efficacy parameters from daily evening pain diary and daily bowel movement diary and PRO data will be presented in subject data listings, for all enrolled subjects. Listings will be ordered by treatment subject number, and assessment week or event date. The treatment group presented in listings will be based on the planned assignment, unless otherwise noted.

Summary tables will be provided for all randomized subjects. All demographic and baseline data will be presented by treatment arm and overall, unless otherwise specified. Efficacy, PRO and safety data will be presented by treatment arm only. In general, continuous variables will be summarized to include the population sample size (N), number of subjects with available data (n), arithmetic mean, SD, median, minimum, Q1, Q3 and maximum values. Categorical variables will be summarized by the population size (N), number of subjects with available data (n), number of subjects in each category, and the percentage of subjects in each category. Unless otherwise noted, the denominator to determine the percentage of subjects in each category will be based on the number of subjects with available data (n) in the analysis set of interest. Selective ordinal data may be summarized using both descriptive statistics and counts and percentages of subjects in each category, as appropriate.

The data analyses will be conducted using the SAS® System (SAS Institute Inc., Cary, NC) version 9.4 or above. The [redacted]' standard operating procedures will be followed for the validation of all SAS programs and outputs.

Formats

Numeric data will be reported at the precision collected on the eCRF. The reported precision from non eCRF sources may be adjusted to a clinically interpretable number or decimal places.

Unscheduled Visits

Unscheduled visits will be assigned to a study visit using the all-inclusive windows defined in Section 4.2.4. However, for by visit summaries, only the planned assessment time points will be summarized. Assessments at unscheduled visits will be included for "any time On-treatment" timepoints and in data listings, as well as algorithms to determine the maximum.

4.2.3. Premature Withdrawal and Missing Data

All data collected in the study database after the subject's early withdrawal of study treatment should be included in summary and analysis.

If any missing data are present in diary data for any reasons, no imputation will be applied to derive responder since the minimum compliance rules described in responder definition in Section 4.3.1 accounted for missing diary entries. No explicit missing data imputation will be performed for change from baseline since a mixed model for repeated measures (MMRM) will be applied to change from baseline analysis. Missing items from the subject reported outcomes (PROs) will be handled according to the respective measure instructions as described in Section 4.3.2.

In general, missing safety data will not be imputed and only observed values will be analyzed. Data of subjects who withdraw after the screening examination or are not treated will be only listed.

If the relationship of an AE record ("Relationship to investigational product" on AE CRF) is missing this AE will be considered "Probably Related" to the study treatment. If the AE intensity is missing every effort should be made to acquire the information from the investigator. "Severe" will be assigned to a missing intensity for reporting purpose

The general imputation rules of partial missing date for both AE and concomitant medication is detailed below:

Dates missing only the day of the month within a year will adhere to the following conventions:

- The missing day of onset date will be set to:
  - First day of the month that the event occurred, if the onset YYYY-MM is after the YYYY-MM of first double-blind treatment
  - The day of the first run-in treatment, if the onset YYYY-MM is the same as YYYY-MM of the first run-in treatment, but different from YYYY-MM of the first double-blind treatment
  - The day of the first double-blind study treatment, if the onset YYYY-MM is the same as YYYY- MM of the first double-blind treatment and the first single-blind treatment
  - The date of informed consent, if the onset YYYY-MM is before the YYYY-MM of the first single-blind treatment.

- The missing day of end date will be set to:
  - The last day of the month of the occurrence. If the subject died in the same month, then set the imputed date as the death date.

Dates missing both the day and month of the year will adhere to the following conventions:

- Missing day of onset date will be set to:
  - January 1 of the year of the onset, if the onset YYYY is after the YYYY of the first double-blind treatment
  - The date of the first run-in treatment, if the onset YYYY is the same as YYYY of the first run-in treatment, but different from YYYY of the first double-blind treatment
  - The date of the first double-blind treatment, if the onset YYYY is the same as YYYY of the first double-blind treatment
  - The date of informed consent, if the onset YYYY is before YYYY of the first run-in treatment

- The missing date of end date will be set to:
  - December 31 of the year of occurrence. If the subject died in the same year, then set the imputed date as the death date.

Missing items from the subject reported outcomes (PROs) will be handled according to the respective measure instructions as described in Section 4.3.2.

4.2.4. Assessment Windows

4.2.4.1. Study Reporting Periods

Based on study design and variables under consideration, study time periods are defined as below.

Table 1: Definition of Study Reporting Periods

Analysis description	Analyzing Study Period	Start Date	End Date
General	Screening	Date of informed consent	Date of first run-in dose
	Run-in	Date of first run-in single-blind dose	Date of last run-in single-blind dose
	Blinded Treatment (Day 1 – Week 12)	Date of first double-blind dose	Date of last double-blind dose + 14
	Follow-up (up to Week 15)	Date of last dose + 14 days	Date of last contact
Adverse events	Screening	Start date ≥ date of informed consent	End date < date of first single-blind run-in dose
	Run-in	Start date ≥ date of first single-blind run-in dose	End data ≤ date of last single-blind run-in dose
	Blinded Treatment	Start date ≥ date of first double-blind dose	End data ≤ date of last double-blind dose + 14 days
	Follow-up	Start date > date of last dose + 14 days	Date of last contact
Concomitant Medication	Prior medication	Start date > date of first double-blind dose	

Analysis description	Analyzing Study Period	Start Date	End Date
	Concomitant mediation	Start date ≥ date of first double-blind dose, or	End date ≤ date of last double-blind dose + 14 days
		Start date > date of first double-blind dose, but ongoing during double-blind treatment period	End date ≤ date of last double-blind dose + 14 days
	Post medication	Start date > date of last double-blind dose +14 days	
The date of last contact is the maximum date of study discontinuation, follow-up contact or the return of study medication. When lost to follow-up, the latest date of assessment/event in the database will be used as the last known date of the subject in study.			

4.2.4.2. Analysis Visiting Window

Analysis windows will be used for eDiary, safety (vitals and labs) as well as quality of life endpoints.

All measurements including the ones from the unscheduled visits that fall within the "Visit Window" will be used in the windowing algorithm to determine the analysis visit. Relative study day will be derived based on Section 4.3.3. Note that the date of the first double-blind dose is Day 1 and the day before the date of the double-blind first dose is Day -1. Thus, there is no Day 0.

For the safety assessments and quality of life endpoints, if there are multiple values for in one window, the visit closest to the nominal day will be selected for assessing endpoints at a particular visit and for by-visit displays. If the visits are equally distant from the nominal day, then the earlier visit will be selected. All values will be stored in analysis datasets.

Check that all records slotted to a post-treatment visit are greater or equal to 15 days after the last dose date. Any post-treatment visit occurring less than 15 days after the last dose date should be re-slotted to an on-treatment visit using the corresponding relative day windows constructed for that parameter.

Table 1. For eDiary, the following analysis visit window slotting will be used to determine weekly assessments:

Analysis window label	Visit Window
Screening	NA
Run-in	NA
Baseline	[-7, -1]

Analysis window label	Visit Window
Week 1	[1, 7]
Week 2	[8, 14]
Week 3	[15, 21]
Week 4	[22, 28]
Week 5	[29, 35]
Week 6	[36, 42]
Week 7	[43, 49]
Week 8	[50, 56]
Week 9	[57, 63]
Week 10	[64, 70]
Week 11	[71, 77]
Week 12	[78, 84]
Follow-Up for Early Withdrawals	>Last dose date + 14
Follow-Up for Completers	>Last dose date + 14

Note: Only calculate the upper bound of last dose date + 14 if the relative study day of the last dose date is ≥ 71 days

Table 2: Analysis Visit Window Slotting for Safety Laboratory and Vitals

Analysis window label	Nominal visit	Nominal day	Visit Window
Screening	Visit 1	NA	NA
Run-in	Visit 2	NA	NA
Baseline[1]	Visit 3	1	NA
Week 2	Visit 4	15	[2, 22]
Week 4	Visit 5	29	[23, 42]
Week 8	Visit 6	57	[43, 70]
Week 12	Visit 7	85	[71,last dose date + 14,]
Follow-Up for Early Withdrawals	NA	NA	>Last dose date + 14
Follow-Up for Completers	NA	NA	>Last dose date + 14

[1] See section 4.3.4 for definition of baseline
Note: Only calculate the upper bound of last dose date + 14 if the relative study day of the last dose date is ≥ 71 days

Table 3: Analysis Visit Window Slotting for ePRO assessments GIS and Additional Pain questions

Analysis window label	Nominal visit	Nominal day	Visit Window
Screening	Visit 1	NA	NA
Run-in	Visit 2	NA	NA
Baseline[1]	Visit 3	1	NA
Week 2	Visit 4	15	[2, 22]
Week 4	Visit 5	29	[23, 42]
Week 8	Visit 6	57	[43, 70]
Week 12	Visit 7	85	[71, last dose date + 14]
Follow-Up for Early Withdrawals	NA	NA	>Last dose date + 14
Follow-Up for Completers	NA	NA	>Last dose date + 14

[1] See section 4.3.4 for definition of baseline

Note: Only calculate the upper bound of last dose date + 14 if the relative study day of the last dose date is ≥ 71 days

[redacted]

Table 5: Analysis Visit Window Slotting for Weekly Pain and Bowel Movement Diary

Analysis window label	Visit Window
Screening	NA
Run-in	NA
Baseline	[-7, 1]
Week 1	[1, 7]
Week 2	[8, 14]
Week 3	[15, 21]
Week 4	[22, 28]
Week 5	[29, 35]
Week 6	[36, 42]
Week 7	[43, 49]

Analysis window label	Visit Window
Week 8	[50, 56]
Week 9	[57, 63]
Week 10	[64, 70]
Week 11	[71, 77]
Week 12	[78, 84]

4.3. Data Definitions and Derivations

4.3.1. Efficacy Endpoints

Efficacy Assessments will be collected as outlined in Table 6 below. Refer to protocol section 1.3 (Schedule of Assessments) for the details on the day of visit and timing of measurement.

Table 6: Efficacy Assessments

Assessment	Form	Measurement and Timing
Abdominal Pain Intensity	Daily Evening Pain Diary (eDiary)	Subjects will record daily worst abdominal pain over the past 24-hours with 0- to 10-point NRS (0 = no pain, 10 = worst possible pain) during both Run-in and Double-blind Period
PRO Questionnaires	• Global Improvement Scale (GIS) [redacted]	• GIS: Day 1, Weeks 2, 4, 8, 12 [redacted]
Additional abdominal pain questions	Daily Evening Pain Diary (eDiary)	Day 1, Weeks 2, 4, 8, 12 • Abdominal pain within 1 hour of eating • Abdominal pain associated with a bowel movement
[redacted]		

Assessment	Form	Measurement and Timing
[redacted]		

There are responders and efficacy endpoints that require further details to clearly define the endpoints without ambiguity. The endpoints and their definitions are described in detail below.

4.3.1.1. API Weekly Responder

To derive API Weekly Responder during various week intervals, the initial two steps should be followed to derive (1) Weekly average worst abdominal pain (WAP) score; (2) Weekly Responder status by week

Step 1:

- Weekly average is defined as the sum of WAP scores on the NRS (0-10) in the 7 days immediately prior to baseline visit (Day 1) or each of post-baseline Weeks (Week 1-Week 12), divided by the number of days with non-missing pain diary entries during the 7 days. Weekly average will be set missing if there are more than 2 missed diary days over the past 7 days.
- API Weekly Responder with ≥ 30% improvement at each week (i.e., Week 1, Week 2, Week 3, …, Week 12) is defined as a subject who experiences a decrease in the weekly average of WAP score of at least 30% compared with the baseline weekly average. To qualify as a responder for the specific week, a subject must have at least 5 days of non-missing pain diary entries during the week (i.e., ≥ 5 of 7 days). A Subject will be assigned as non-responder for the specific week if the weekly average is missing.

Similarly, the API Weekly Responder will be further derived based on more stringent measures of weekly reduction as below:

- API Weekly Responder with ≥ 40% improvement at each week (i.e., Week 1, Week 2, Week 3, …, Week 12) is defined as a subject who experiences a decrease in the weekly

average of WAP score of at least 40% compared with the baseline weekly average. To qualify as a responder for the specific week, a subject must have at least of 5 days of non-missing pain diary entries during the week (i.e., ≥ 5 of 7 days).

- API Weekly Responder with ≥ 50% improvement at each week (i.e., Week 1, Week 2, Week 3, …, Week 12) is defined as a subject who experiences a decrease in the weekly average of WAP score of at least 50% compared with the baseline weekly average. To qualify as a responder for the specific week, a subject must have at least of 5 days of non-missing pain diary entries during the week (i.e., ≥ 5 of 7 days).

Step 2:
Then API Weekly Responder will be derived during different treatment intervals using primary criteria as an example below:

<u>API Weekly responder with ≥ 30% improvement over 12 weeks (i.e., from randomization through to Week 12) - primary efficacy endpoint</u>

- To qualify as an overall API Weekly Responder from randomization through to Week 12 (Weeks 1-12), a subject must satisfy the following criteria:
  - Must meet the API Weekly Response criteria ≥ 30% improvement on ≥ 50% of weeks, i.e., success on 6 or more weeks during the past 12 weeks
  - Must have had ≥ 60 diary days over the past 12 weeks

<u>API Weekly responder from randomization through to Week 2 (Weeks 1-2)</u>

- To qualify as an API Weekly Responder from randomization through to Week 2, a subject must satisfy the following criteria:
  - Must meet the API Weekly Response criteria on more than 50% of weeks, i.e., success on 1 or 2 weeks during past 2 weeks
  - Must have had ≥ 10 diary days over the past 2 weeks

<u>API Weekly Responder from randomization through to Week 4 (Weeks 1-4)</u>

- To qualify as an API Weekly Responder over 4-week interval, a subject must satisfy the following criteria:
  - Must meet the API Weekly Response criteria on more than 50% of weeks, i.e., success on 2 or more weeks during past 4 weeks
  - Must have had ≥ 20 diary days over the past 4 weeks

<u>API Weekly Responder from randomization through to Week 8 (Weeks 1-8)</u>

- To qualify as an API Weekly Responder over 8-week interval, a subject must satisfy the following criteria:
  - Must meet the API Weekly Response criteria on more than 50% of weeks, i.e., success on 4 or more weeks during past 8 weeks
  - Must have had ≥ 40 diary days over the past 8 weeks

The secondary efficacy endpoints of Weekly API Responder with either ≥ 40% or ≥ 50% improvement will be classified using the same rules above.

4.3.1.2. GIS Responder

GIS responder at Week 12 is the secondary efficacy endpoint. It is defined as a subject who selected either 1) (significantly relieved) or 2) (moderately relieved) on the GIS questionnaire at Week 12. A subject with a missing response will be considered as non-responder in the specific visit. GIS responder will be derived at Day 1, Week 2, Week 4, Week 8, and Week 12 for all IBS subjects.

4.3.1.3. Change from Baseline in API score

All endpoints will be derived at Baseline, Weeks 2, 4, 8 and 12.

- Weekly average of WAP score:

  Derivation is defined as described in the first step in Section 4.3.1.1.

- Weekly average number of pain-free days

  Weekly average number of pain-free days is defined as the sum of all pain-free days in the past 7 days immediately prior to baseline visit (Day 1) or each of post-baseline Weeks (Weeks 2, 4, 8, 12), divided by the number of days with non-missing pain diary entries and then multiply 7. Weekly average will be set missing if more than 2 missed diary days over the past 7 days.

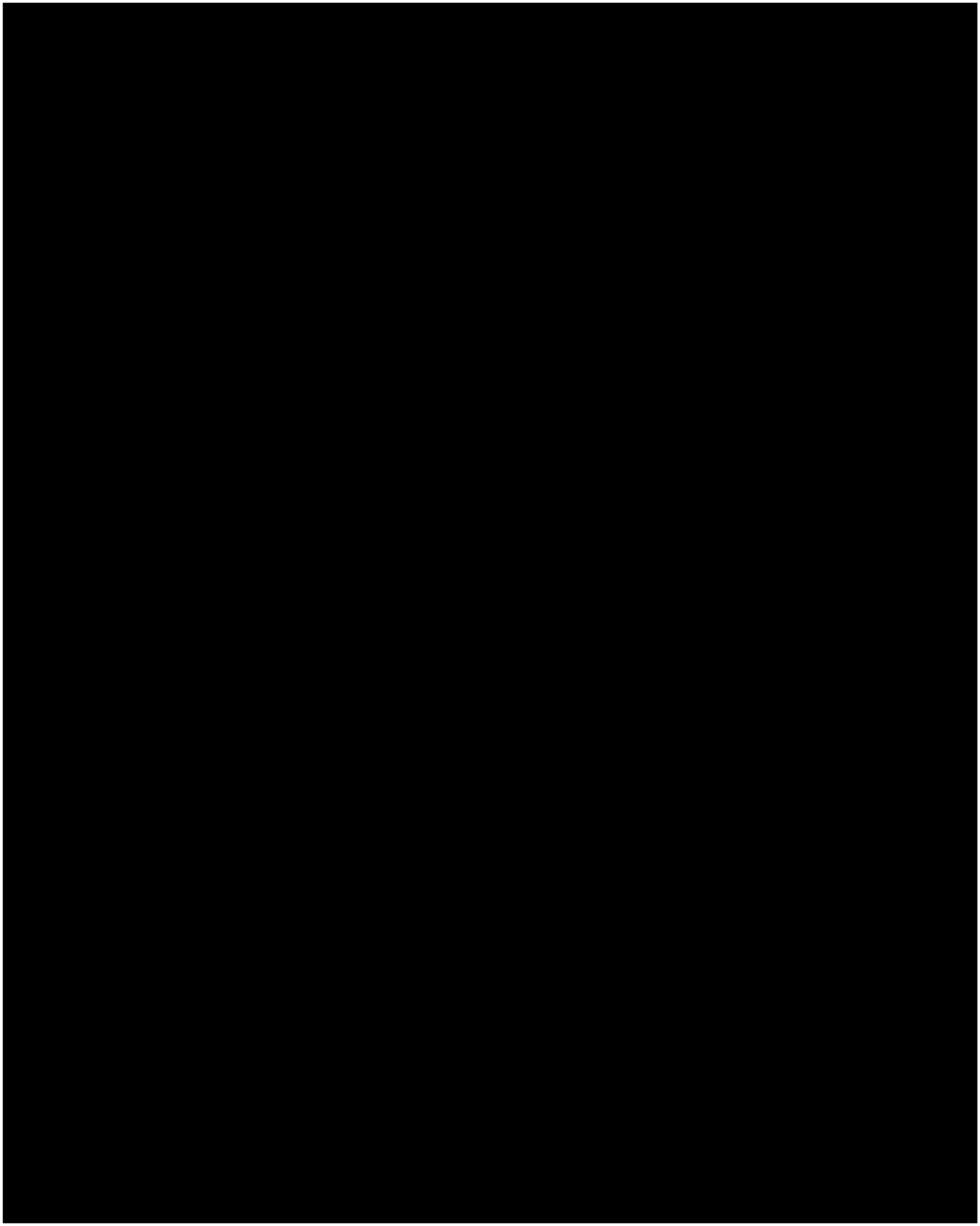
- Change from baseline in abdominal pain within 1 hour of eating

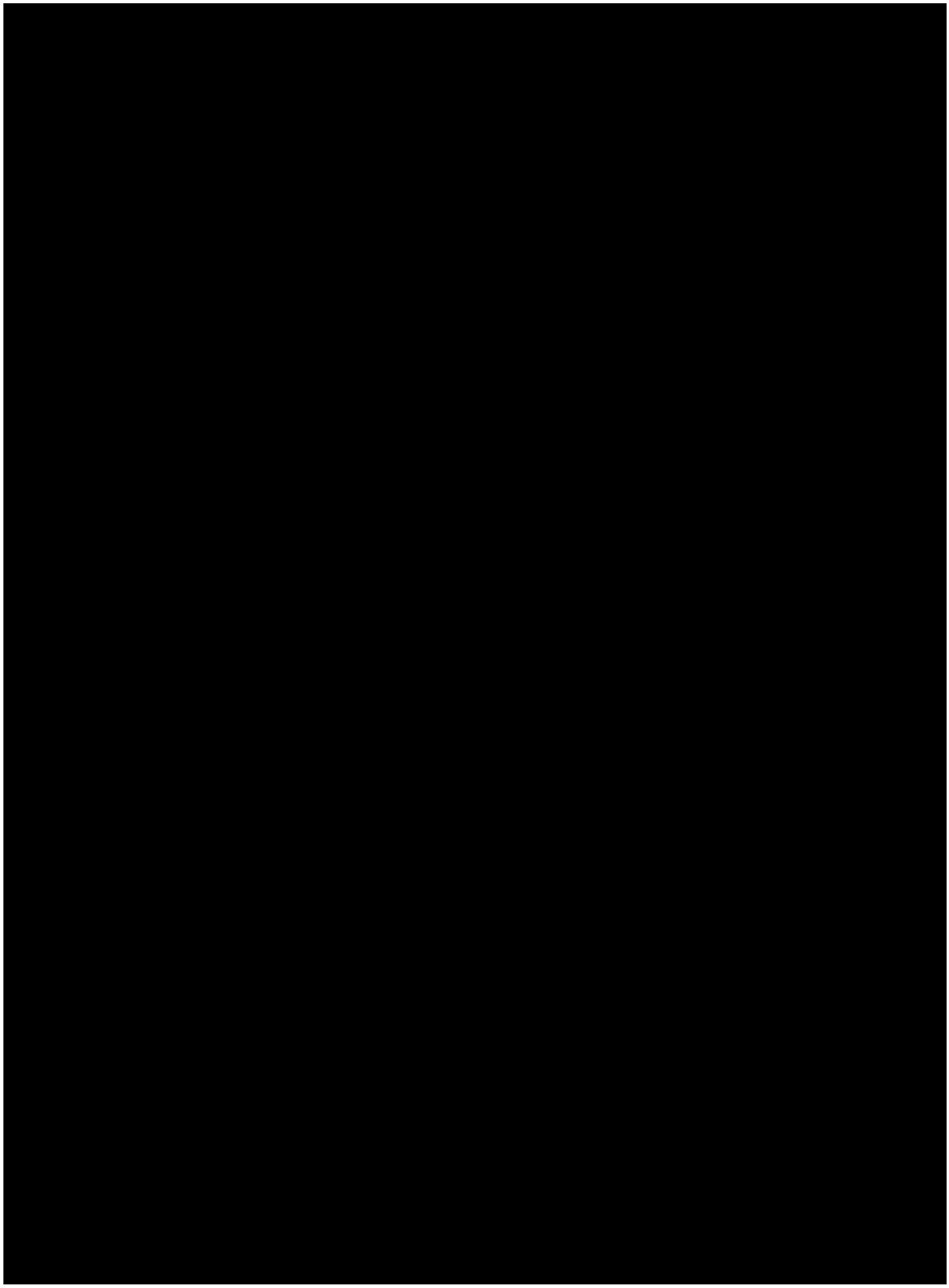
  Change from baseline will be summarized categorically as “Improved (Yes to No)” vs “Not Improved”, where “Not Improved” is further categorized as “No Change”, “Worsened (No to Yes)”, or “Missing Post-baseline Assessment”.

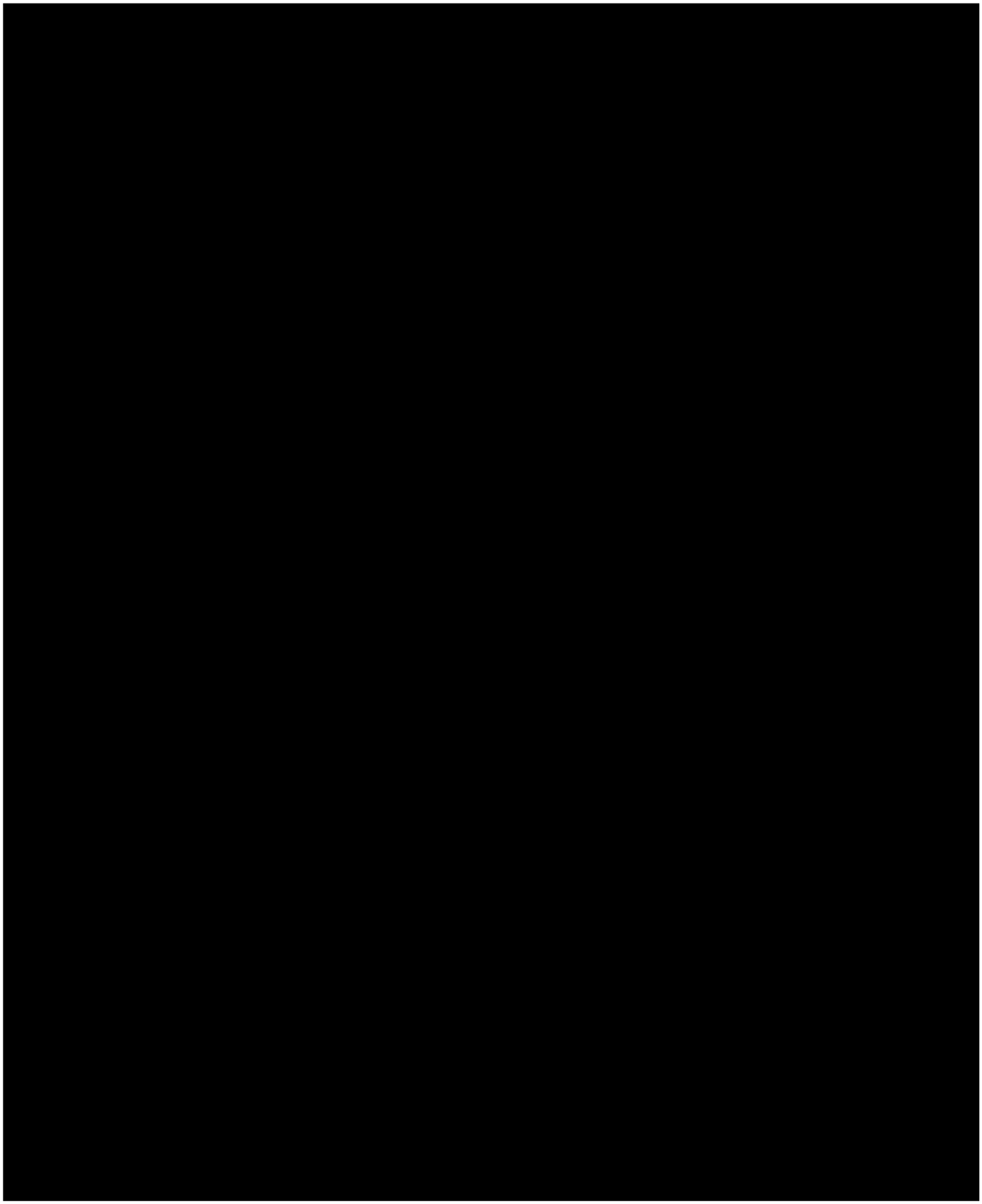
- Change from baseline in abdominal pain associated with a bowel movement:

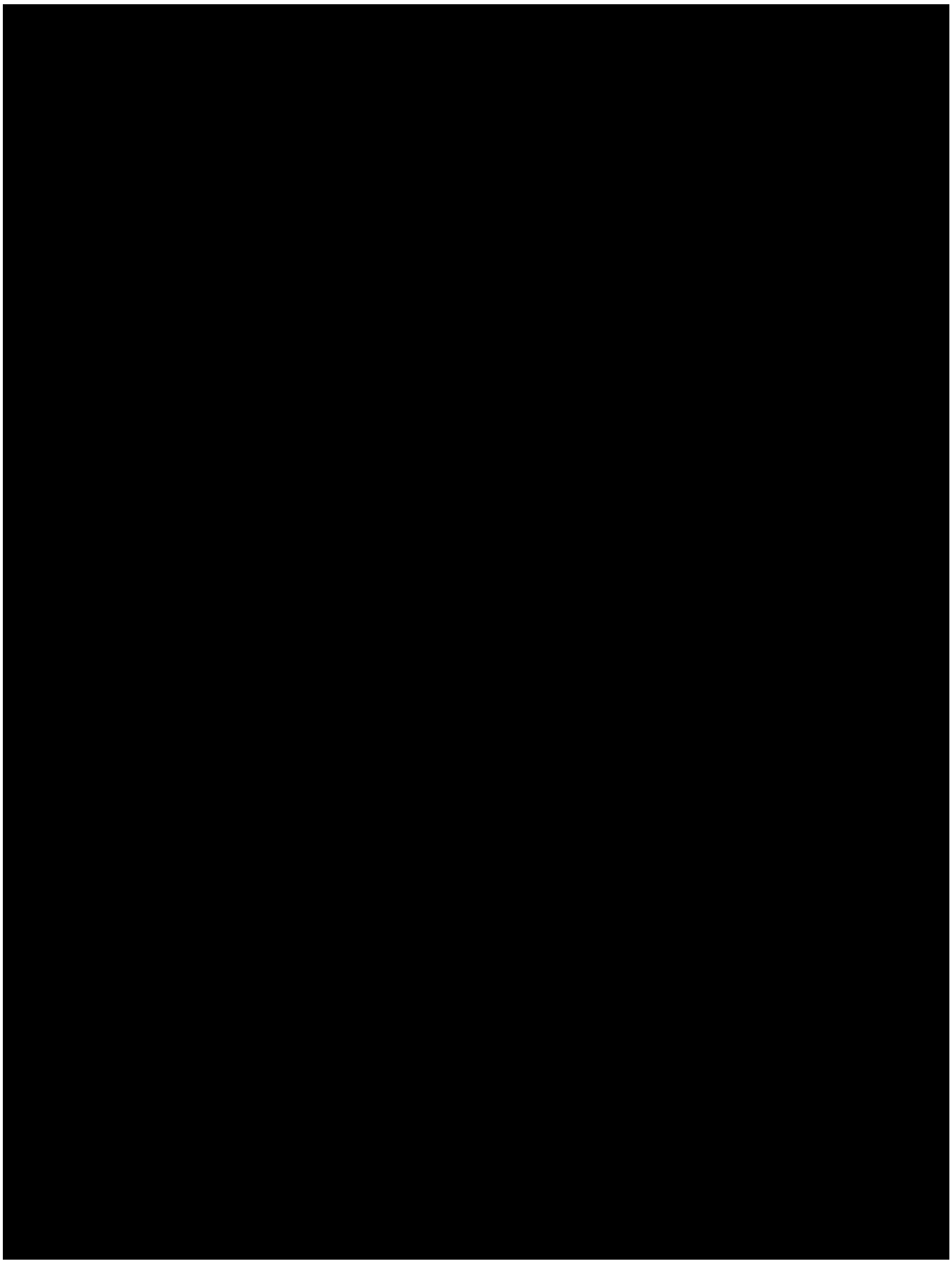
  Change from baseline will be summarized categorically as “Improved (Yes to No)” vs “Not Improved”, where “Not Improved” is further categorized as “No Change”, “Worsened (No to Yes)”, or “Missing Post-baseline Assessment”.

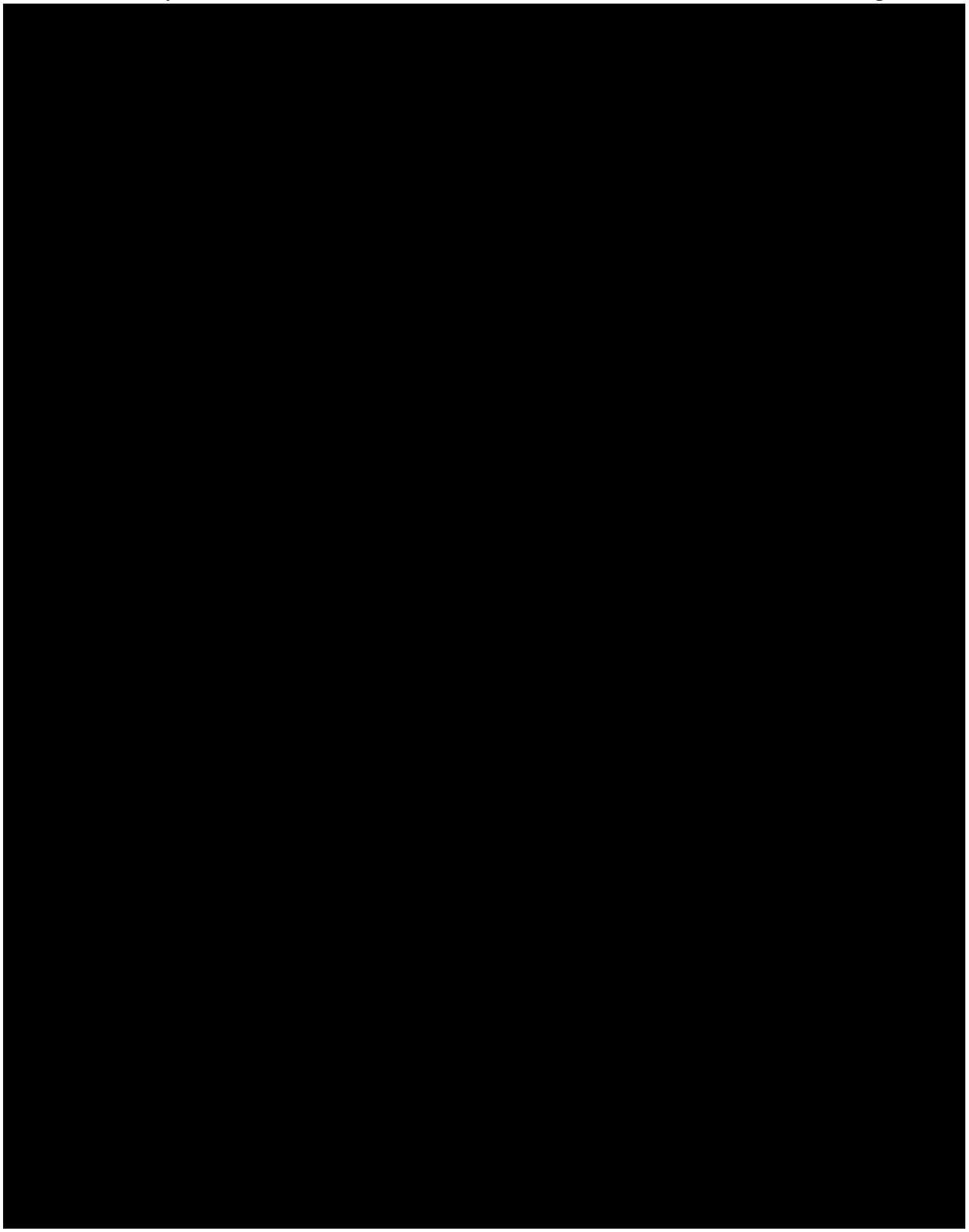
[redacted]

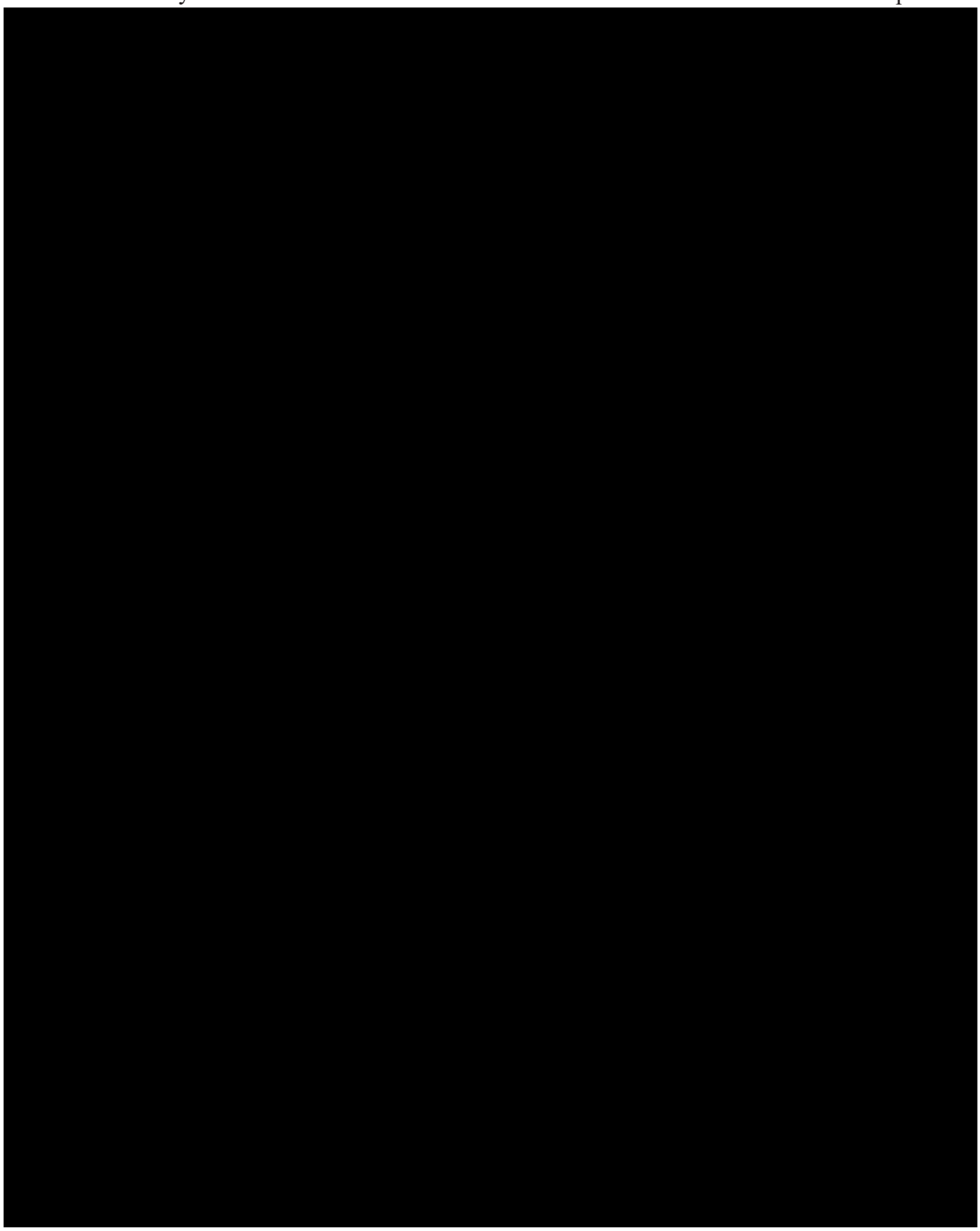




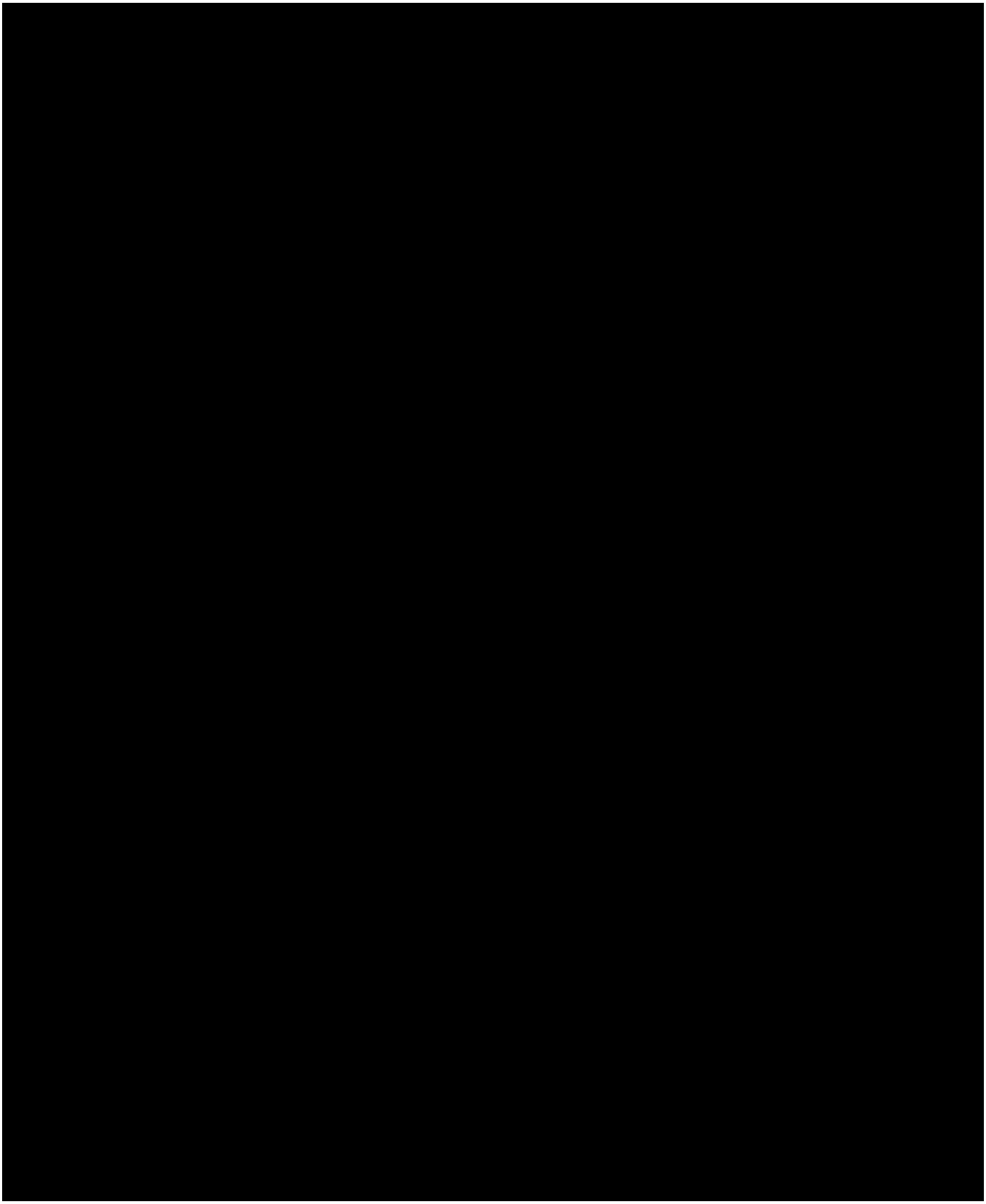








4.3.2. Reporting Questionnaire Scoring



4.3.3. Study Day and Duration

Study day is relative to the start date of the double-blind treatment. This is used to describe the relative time of an event or assessment happened during the study. The first day of the study is defined as the day a subject first receives either vibegron or placebo in the double-blind treatment period. This is expected to be the same day as the randomization as Day 1. There is no study Day 0 is defined in the study.

- For an event or assessment that occurred on or after the first dose of double-blind treatment date:

:

Study day = Date of event or examination – date of first double-blind study treatment + 1

- For an event or assessment that occurred prior to the first dose of double-blind treatment date:
  Study day = Date of event or examination – date of first double-blind study treatment

Similarly, a duration between any two dates (such as AE duration) expressed in days will be calculated using the following conventions:

- Duration = Later date – earlier date + 1, if the earlier date is on or after the date of first dose of double-blind treatment

- Duration = Later date – earlier date, if the earlier date is prior to the date of first dose of double-blind treatment

4.3.4. Baseline and Change from Baseline

In general, the last recorded value on or prior to the date of double-blind treatment start will serve as the baseline measurement for efficacy endpoints while the last recorded value prior to first dose of double-blind treatment will serve as the baseline measurement for safety endpoints. The mean of the multiple values will be used as baseline for the following situations:

- If multiple measurements are scheduled on the same baseline day (i.e., blood pressure)
- If multiple measurements are collected on the same baseline day without the time or "repeat" status to differentiate the records

Change from baseline will be calculated as the post-baseline value minus the baseline value. Percentage change from baseline will be calculated as the change from baseline divided by the baseline value, multiplied by 100. If either the baseline or post-baseline value is missing, then change from baseline and percentage change from baseline will be set to missing.

5. DEMOGRAPHIC, OTHER BASELINE CHARACTERISTICS AND MEDICATION

5.1. Subject Disposition and Withdrawals

Subject disposition will be summarized by treatment group for the Screened, Run-in and Randomized Sets separately. The summary table will present the frequency and percentage of subjects in each of the analysis sets and those who discontinued the study prematurely along with the primary reasons for discontinuation.

For the summary under the Randomized Set, the following additional categories will be presented as well; randomized, received treatment with study drug, did not receive treatment with study drug, completed treatment with study drug, discontinued treatment with study drug (and reason), and withdrawn from study (and reason).

The frequency and percentage of subjects with at least one major Protocol Deviation (PD), major PD by classification and reasons/category for PD will be summarized by treatment arm and by FAS-D, FAS-M and FAS. Inclusion in each of the analysis sets (SAF, FAS, FAS-D, etc.), and any reasons for exclusion will be summarized by treatment arm for the Randomized Set.

Screen Failure, Run-in and Double-Blind Period disposition, with reasons for discontinuation of study will also be listed, including the date of discontinuation and date of last contact. The date of last contact is the maximum date of study discontinuation, follow-up or the return of study medication. When lost to follow-up, the latest date of assessment/event in the database will be used as the last known date of the subject in study.

Eligibility criteria, screening and run-in failures (including date and primary reason for failure), and informed consent (protocol version, informed consent version date and date signed) will be listed for all subjects screened.

A summary of randomized subjects by country, investigator name and site identification will be provided.

Randomization details will also be listed, including the date of randomization, randomization number and randomization strata. A listing will also be provided for the subjects who had misclassification of the stratification factors (i.e., randomization stratification does not match actual stratification values). The actual stratification is based on medical history lower level terms of "Irritable bowel syndrome" which indicates the subjects has a history of IBS-M and the term of "Diarrhea predominant irritable bowel syndrome" which indicates the subject has a history of IBS-D. If more than one term is available, the term ongoing just prior to randomization should be used. In the rare event that no medical history term is able to be mapped to either IBS-D or IBS-M then the actual stratification will be assigned the value of the randomization strata.

5.2. Demographic and Baseline Characteristics

All demographic and baseline characteristic data will be summarized by treatment group using descriptive statistics for all IBS subjects, IBS-D and IBS-M based on the FAS. In addition, the summary will be prepared for PPS-D.

The summary table will include age, age category ((<40, ≥40 to <65, ≥65 years), sex, child bearing potential (Yes/No), ethnicity, race, baseline abdominal pain intensity score (<6 vs ≥ 6), baseline IBS subtype (IBS-D vs. IBS-M) (only for all IBS subjects), previous IBS diet (Yes/No), category of previous IBS diet if Yes, baseline hypertension category (Hypertension vs. normotensive), pre-existing hypertension (Yes/No), weight, height and BMI.

Baseline Hypertension will be defined as baseline systolic blood pressure (SBP) ≥140 mmHg or baseline diastolic blood pressure (DBP) ≥90 mmHg, regardless of medical history.

Pre-existing hypertension will be defined as having a medical history of hypertension or baseline hypertension (baseline SBP ≥140 mmHg or baseline DBP ≥90 mmHg). The following list of preferred terms will be used to search for medical history of hypertension:

- Accelerated hypertension
- Diastolic hypertension
- Essential hypertension
- Hypertension
- Hypertensive crisis
- Hypertensive emergency
- Hypertensive heart disease
- Malignant hypertension
- Malignant hypertensive heart disease

- Secondary hypertension
- Supine hypertension
- Systolic hypertension

Age (years), height (cm), weight (kg), and BMI captured at Screening will be summarized as a continuous variable.

All demographic data will be listed.

5.3. Other Baseline Characteristics

The data from the Daily Evening Pain Diary and Bowel Movement diary during run-in period (reviewed at Visit 3 Baseline visit) prior to first double-blind dose will be used as baseline for each subject. This includes endpoints:

- Weekly average of worst daily pain score,
- Weekly average number of pain-free days
- Daily average bowel urgency episodes,
- Weekly average number of days with bowel urgency,
- Abdominal pain within 1 hour of eating (Yes vs No),
- Abdominal pain associated with a bowel movement, (Yes vs No)
- Weekly average number of days with recurrent bowel movement,
- Daily average number of bowel movements,
- Weekly average number of days with diarrhea (Bristol Type 6 or 7),
- Weekly number of days with diarrhea (Bristol Type 6 or 7)
- Weekly average number of days with at least 1 stool with Bristol Type 3, 4, or 5 consistency
- Weekly average number of stools with Bristol Type 3, 4, or 5 consistency

In addition, the following parameters will be presented in the baseline table

- Daily average stool consistency
- Daily average number of diarrhea bowel movement
- Daily average number of stools with Bristol Type 3, 4, 5 consistency
- Weekly average number of days with constipation (Bristol Type 1 or 2)
- Daily average number of stools with constipation (Bristol Type 1 or 2)
- Weekly average number of stools with constipation (Bristol Type 1 or 2)

These will be summarized by IBS subtype (IBS-D, IBS-M) and for all IBS subjects, and treatment group using descriptive statistics for continuous data for all subjects in each analysis set.

5.4. Medical History and Concomitant Disease

Descriptions of medical history findings will be coded using Medical Dictionary for Regulatory Activities (MedDRA) version 23.0 or higher.

A disease or illness reported as medical history without a start date will be included in medical history without a date assigned. Medical history will be sorted by descending overall frequency, by system organ class (SOC) and preferred term (PT) in the summary table. Medical history data listings will be sorted by treatment, subject number, start date, SOC and PT.

5.5. Prior and Concomitant Medication

Prior and concomitant medications will be coded using the World Health Organization (WHO) Drug B3-March Format, 2020 version or later. These medications will be further classified as follows.

- Prior medication: are defined as those medications taken prior to the first dose of double-blind treatment.
    - Non-IBS prior medication
    - IBS prior medication
- Concomitant medication: are defined as those taken on or after the first dose of double-blind medication but prior to the last dose of double-blind medication + 14 days. If medications started prior to double-blind medication and were ongoing during the Double-Blind Period, those medication will be considered as concomitant medication as well. See Table 1.
- Post medication: medication taken after last dose of double-blind medication + 14 days

Except for Prior IBS medication, the number and percentage of subjects taking prior non-IBS medications and non-rescue concomitant medications will be summarized overall by ATC (Anatomical Therapeutical Chemical) Levels 2 and 4 for all subjects in the SAF. Prior IBS medications and rescue concomitant medications will be summarized by ATC Levels 2, 4 and Preferred Term in the SAF. Prior medications, IBS rescue concomitant and non-rescue concomitant medications will be listed for all subjects in the SAF. Post medication if any will be listed only.

5.5.1. Rescue Medication

Use of rescue medication will be collected on the daily diary (i.e., subjects will respond to prompts asking if any rescue medication was taken for abdominal pain or stool or pain/stool symptoms). The proportion of subjects who used rescue medication and the reason for the medication will be summarized by treatment groups. This data will be summarized for the efficacy populations.

Note, protocol Amendment 3 clarified that rescue medications should only be used beginning at Day 1 (Visit 3).

6. EFFICACY ANALYSIS

In general, the FAS will be used for all analyses of efficacy endpoints unless stated otherwise.

6.1. Primary Efficacy Analyses

6.1.1. Primary Analysis

The primary efficacy endpoint of proportion of IBS-D subjects who are API weekly responders with ≥30% improvement from Week 1 through Week 12 is defined in the Section 4.3.1.

The number and percentage of responders and non-responders will be tabulated by treatment and visit. The treatment comparison will use a Cochran-Mantel-Haenszel (CMH) common risk difference estimate stratified by baseline abdominal pain (< 6 vs. ≥6) strata per randomization stratification [11] with weights proposed by Greenland and Robins [6], which is calculated as follows:

$\hat{\delta}_{MH} = \frac{\sum_{i=1}^{u} w_i \cdot \hat{\delta}_i}{\sum_{i=1}^{u} w_i}$, where

$\hat{\delta}_i = \frac{x_i}{n_i} - \frac{y_i}{m_i}$ denotes the risk difference in stratum $i, i = 1, \cdots, u$

$w_i = \frac{n_i \cdot m_i}{n_i + m_i}$ denotes the weight of stratum $i, i = 1, \cdots, u$

x_i denotes the number of subjects with event in treatment_1 in stratum i, $i = 1, \cdots, u$

y_i denotes the number of subjects with event in treatment_2 in stratum i, $i = 1, \cdots, u$

n_i denotes the number of subjects on treatment_1 in stratum i, $i = 1, \cdots, u$

m_i denotes the number of subjects on treatment_2 in stratum i, $i = 1, \cdots, u$

The estimated variance of $\hat{\delta}_{MH}$ is calculated as:

$$\widehat{var}\left(\hat{\delta}_{MH}\right) = \frac{\sum_{i=1}^{u} L_i}{\left(\sum_{i=1}^{u} w_i\right)^2}$$

where $L_i = \frac{x_i(n_i - x_i)\, m_i^3 + y_i(m_i - y_i)\, n_i^3}{n_i \cdot m_i \cdot (n_i + m_i)^2}$, $i = 1, \cdots, u$

Assuming a normal distribution of $\hat{\delta}_{MH}$, an approximate 90% CI is given as follows, where $z_{0.95}$ is the 95% quantile of the standard normal distribution:

$$CI = \left[\hat{\delta}_{MH} \pm z_{0.95} \cdot \sqrt{\widehat{var}\left(\hat{\delta}_{MH}\right)}\right]$$

Also, the approximate p-value can be calculated using the following:

$$\text{p} - \text{value} = 2 \cdot \Pr\left[Z > \left|\frac{\hat{\delta}_{MH}}{\sqrt{\widehat{var}\left(\hat{\delta}_{MH}\right)}}\right|\right], \text{where } Z \sim N(0,1)$$

If there is a stratum for a treatment group that has 0 subjects in it, the 0 count will be replaced by 0.5 in order to prevent dividing by 0 in the above equations, as suggested in Greenland and Robins.

The estimated common risk difference, and associated p-value and 2-sided 90% confidence interval will be tabulated. The stratified strata is based on the baseline abdominal pain score which is the average from 7 days immediately prior to the baseline visit. If more than 10% mis-classification of the stratification factors (randomization stratification vs. actual stratification values) happens a sensitivity analysis will be carried out to repeat the primary analysis where the actual stratification will be used in the model (See Section 6.1.2).

The number and percentage of responders with ≥30% improvement over Weeks 1-2, Weeks 1-4, Weeks 1-8 and Weeks 1-12 will be summarized by treatment. The proportion of subjects who meet the Weekly responder criteria at each week will be plotted by treatment.

To assess the impact of missing data, a categorical summary of missing pain and bowel movement diary day entries, summarized by week, will be provided with the categories of none, 1-2, > 2 days, >3. The denominator used for the percentages will be the number of subjects who completed diaries for the given week.

6.1.2. Sensitivity Analysis

The following sensitivity analyses will be performed:

6.1.2.1 **API weekly responder ≥30% improvement from Week 1 through Week 12 using PPS-D**

The treatment comparison of proportion of IBS-D subjects who are API weekly responders with ≥30% improvement from Week 1 through Week 12 will be performed using CMH described in the primary analysis based on the PPS.

6.1.2.2 **API weekly responder ≥30% improvement from Week 1 through Week 12 using Logistic Regression**

In addition, Logistic regression will be used as a sensitivity analysis to analyze the API weekly responder with ≥30% improvement over Weeks 1-12 based on the FAS but will be considered

exploratory. Responder proportion over Weeks 1-12 will be modeled using treatment and baseline abdominal pain (< 6 vs. ≥6) strata per randomization stratification fixed effects. A summary table will present the estimated incidence for each treatment along with the odds ratio, and associated corresponding 90% CI and p-value for the treatment comparison. Wald statistics and Wald confidence limits from SAS PROC LOGISTIC will be used for the test and the construction of 90% CI.

An example of the SAS code for the base procedure is given below if responder is assigned to 1 and no-responder is assigned to 0.

6.1.2.3 **API weekly responder accounting for the use of rescue medication**

Additionally, a sensitivity analysis of the primary analysis will be performed to assess the impact of subjects receiving rescue medications for pain where subject's diary will be considered non-evaluable on the days that they received rescue medications. This analysis approach will also be performed in FAS-D to assess the treatment comparison of API weekly responders ≥40% and ≥50% improvement, respectively.

6.1.2.4 **API weekly responder based on actual randomization**

The stratified strata is based on the baseline abdominal pain score which is the average from 7 days immediately prior to the baseline visit. If more than 10% mis-classification of the stratification factors (randomization stratification vs. actual stratification values) happens a sensitivity analysis of the primary analysis will be carried out to repeat the primary analysis where the actual stratification will be used in the model.

6.1.3. Subgroup Analysis

The CMH estimate of treatment difference for API weekly responder with ≥30% improvement over Weeks 1-12 with will be prepared for each of stratified strata (abdominal pain intensity score <6 vs ≥ 6) based on the FAS.

In addition, the proportion of responder will be summarized by the following three subgroups.

- Age category (<40, ≥40 to <65, ≥65 years)
- Race (as recorded in CRF)
- Subjects randomized prior to 31Dec2019 vs Subjects randomized after 31Dec2019
- Baseline Bristol Stool Scale (< 5.5 vs ≥ 5.5)

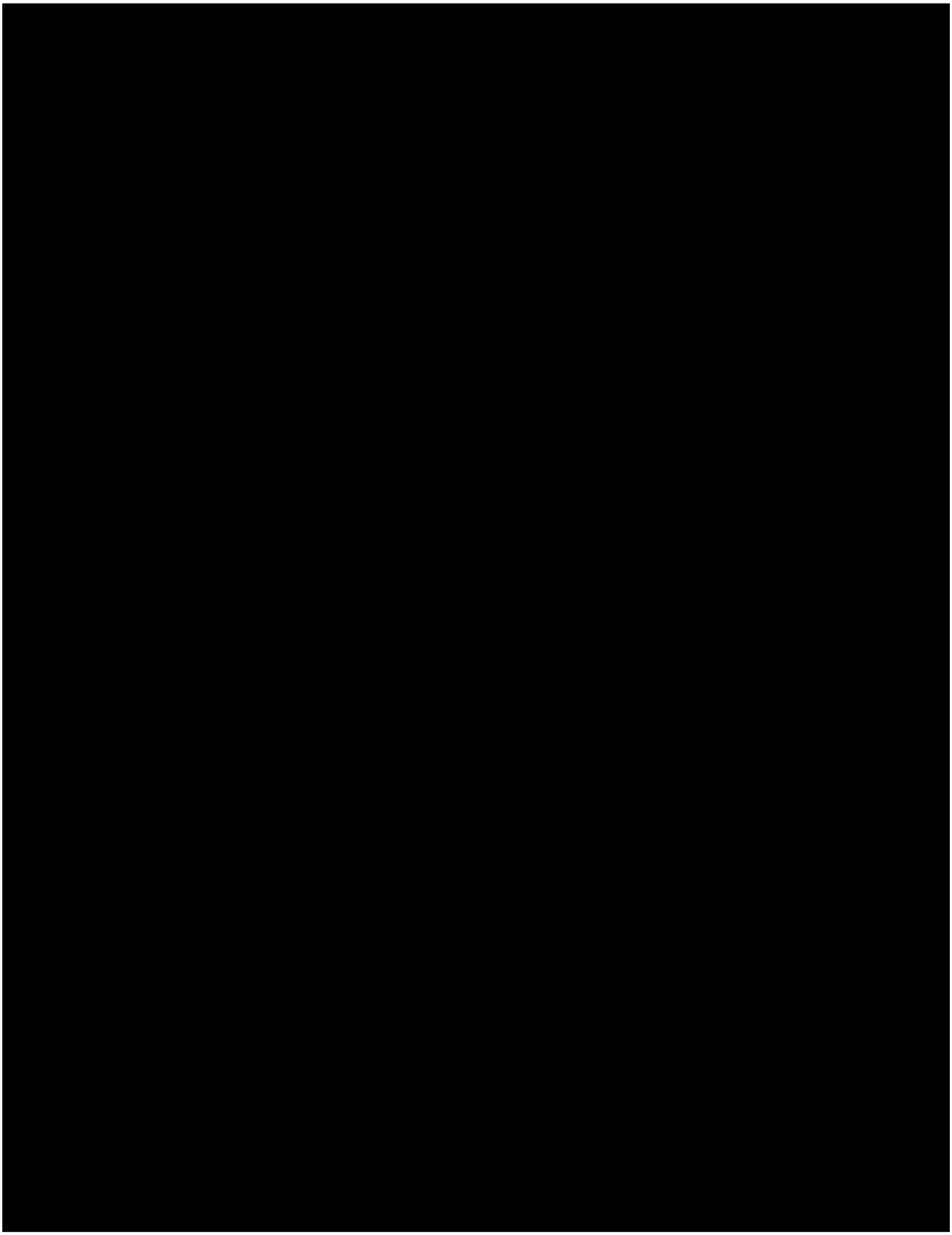
6.2. Secondary Efficacy Analyses

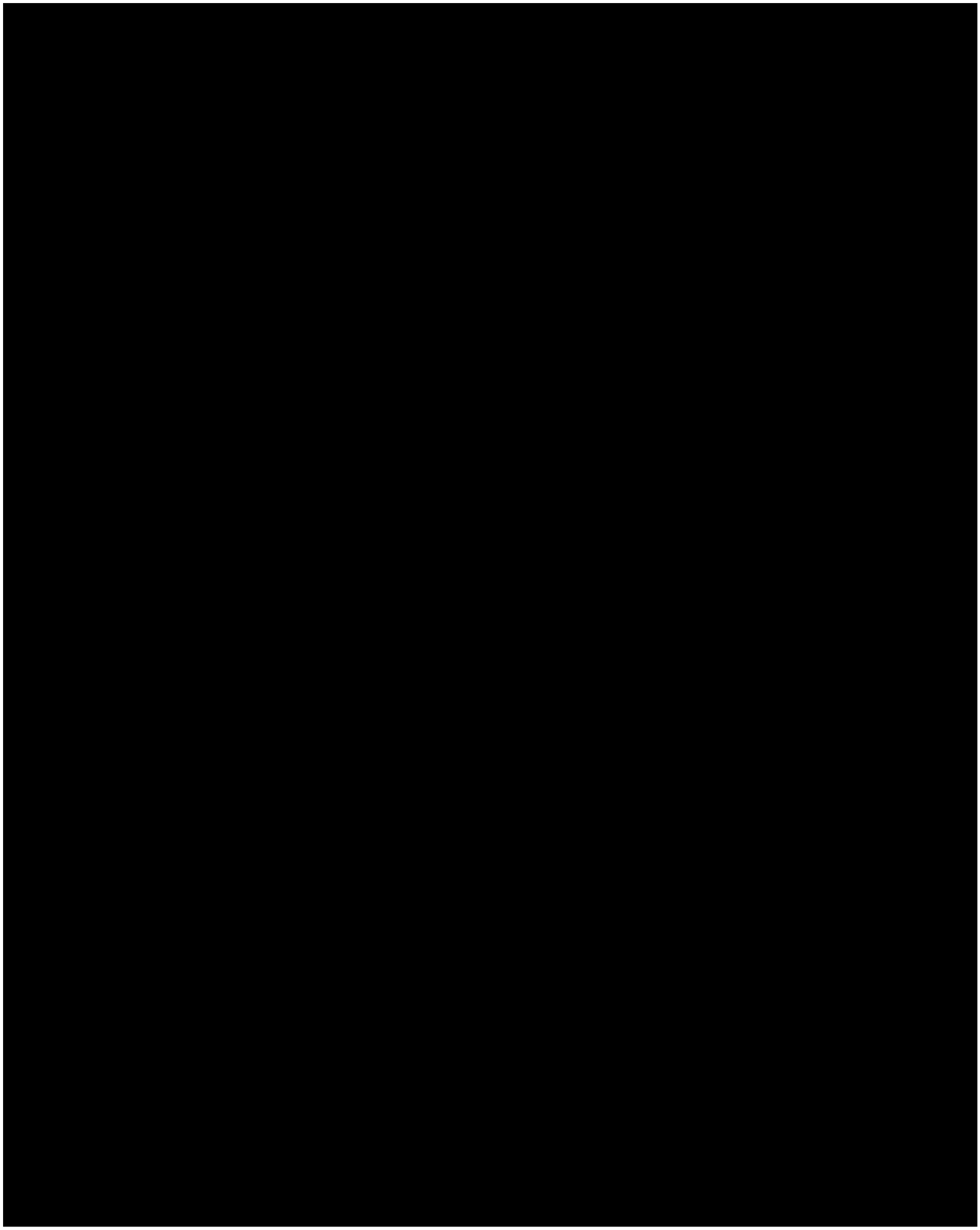
The secondary efficacy endpoint GIS responder at Week 12 is defined in Section 4.3.1. The proportion of responders will be analyzed similarly as described for primary endpoints using CMH for IBS-D and IBS_M subjects. When all subjects are included for the analysis, additional IBS subtype (IBS-D vs. IBSM will be used for stratification as well. The number and percentage of responder and non-responder for IBS-D, IBS-M and all IBS subjects will be tabulated by treatment and visit. GIS response data will be listed.

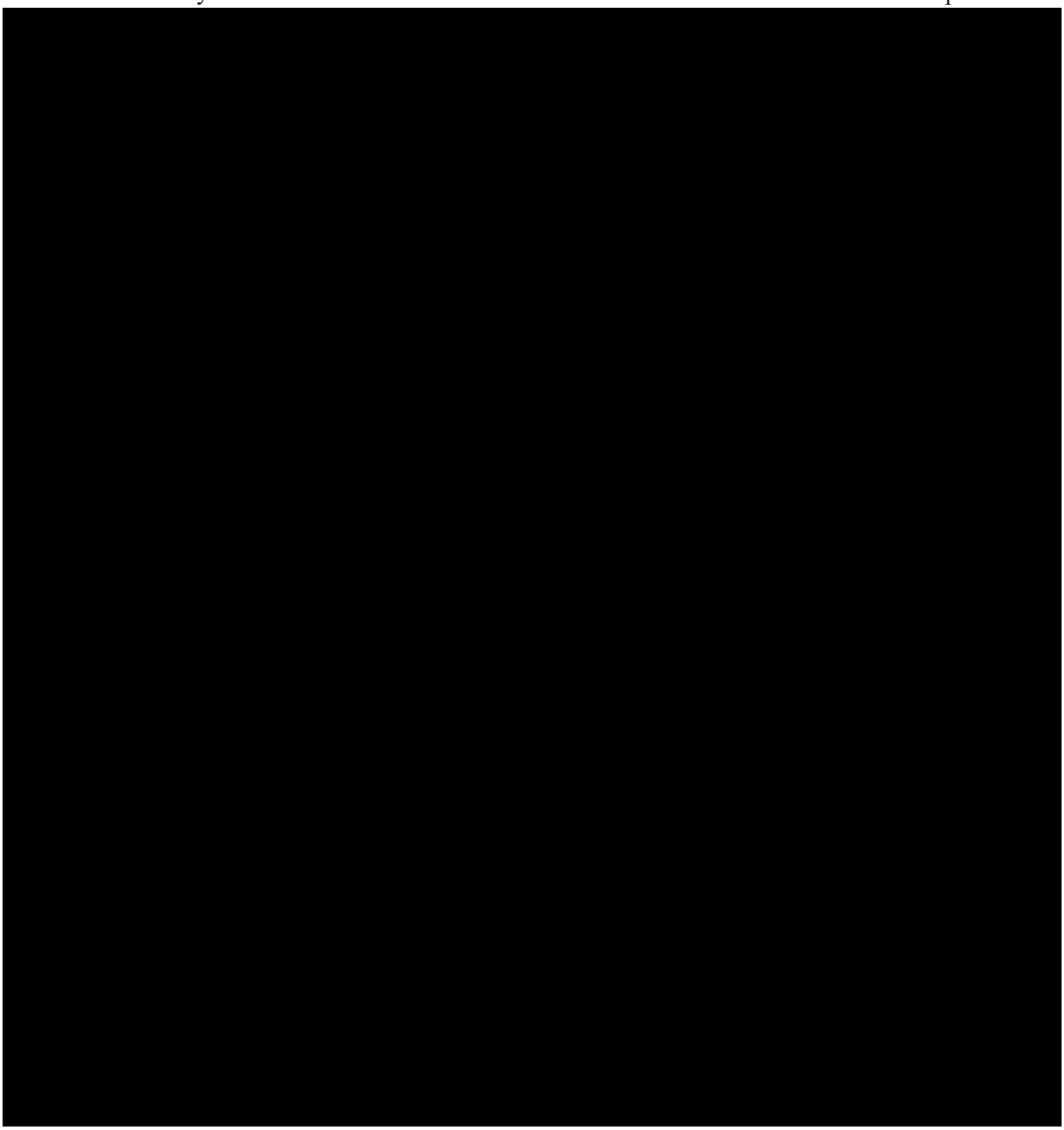
The secondary efficacy endpoints of proportion of IBS-D subjects who are API weekly responders with 40% improvement and API weekly responders with 50% improvement from Week 1 through Week 12 are defined in the Section 4.3.1. The similar statistical analyses described in primary efficacy endpoint will be performed for these API weekly responders including subgroup analysis.

The secondary efficacy analyses will be primarily based on the FAS, unless stated otherwise.

6.3. Exploratory Efficacy Analysis







7. SAFETY ANALYSIS

The SAF will be used for all safety analyses. Safety will be assessed on the basis of extent of exposure and compliance, AE reports, clinical laboratory data, ECGs, physical examinations, and vital signs.

No inferential statistical testing is planned on the safety data, all data will be descriptively summarized by treatment and visit.

7.1. Extent of Exposure

The duration of exposure during the double-blind treatment period will be expressed as the time in days from the first dose as recorded on the Study Drug Administration CRF page at Visit 3 through to last treatment day (inclusive) as recorded on the End of Treatment CRF

page, excluding any days where it is recorded that an interruptions has occurred on the AE CRF page and as indicated on the COVID19 CRF page.

Interruption in days = date started – date stopped + 1

This is given by the following formula:

$$Duration\ (days) = date\ last\ double\ blind\ dose - date\ first\ double\ blind\ dose - interruption\ days + 1$$

Duration of exposure will be summarized by treatment group for the SAF using summary statistics for continuous variables. The proportion of subjects with Dose Interruptions ≥ 2 weeks (14 consecutive days) will be summarized by treatment group, if applicable.

A listing will present the treatment start and end date together with the date of interruption, and the overall days of exposure.

7.2. Treatment Compliance

Study treatment compliance (%) will be calculated as the actual number of doses divided by the expected number of doses, multiplied by 100 and summarized by treatment group for double-blind treatment period

These numbers will be determined by the number of tablets dispensed and returned unused by the subject. Where no treatment bottle is returned, and thus the actual number of doses is unknown, it will be assumed that the subject took all medication available in the bottle. The overall compliance will be calculated for double-blind treatment period, respectively.

$$Overall\ Compliance\ (\%) = \frac{Number\ of\ total\ tablets\ taken}{Target\ number\ of\ total\ tablets\ dispensed} \times 100\%$$

The number of tablets taken is obtained by: (no. dispensed – no. returned). The target number of tablets to be taken is calculated as: (last dose date – first dose date – drug interruption in days + 1). Partial dates will be imputed if the day of the month is missing as described in section 4.2.3. However, if month and/or year are missing for the date of last dose then date of last dose will be assigned the study discontinuation date. Treatment Compliance will be summarized by treatment group and overall for the SAF population. Additionally, the number and percentage of subjects within each treatment with compliance in the following categories will be provided: <80%, 80 – 120%, and >120%. A listing will present compliance data, including bottle number of tablets dispensed, returned, and dispensed and returned dates and overall compliance.

7.3. Adverse Events

AEs will be coded using MedDRA version 23.0. or later.

All reported AEs (whether treatment emergent or not) will be included in by-subject AE listings. Sorting will be by treatment subject, date of event, SOC, PT and then verbatim description.

An AE will be considered treatment emergent (TEAE) if it begins or worsens in severity after the first dose of the double-blind study treatment through 14 days after the last dose of study treatment. Partial AE start dates will be imputed as detailed in Section 4.2.2

Summary tables will be based on TEAEs. The incidence of TEAEs will be presented using counts and percentages of subjects with TEAEs and tabulated by SOC and PT. SOC will be sorted in descending frequency and PT within SOC will be sorted by descending frequency based on the incidence across subjects overall. If a subject has multiple occurrences (start and stop) of an event associated with a specific SOC or PT within a SOC, a subject will only be counted once in the incidence count for the SOC or PT within SOC respectively.

An overall summary table of AEs by treatment group will be presented detailing the number and percentage of subjects, and number of events for the following categories:

- At least one TEAE;
- At least one Treatment-Related TEAE;
- At least one Grade ≥ 3 TEAE (Mild = Grade 1, Moderate = Grade 2, Severe = Grade 3,

Life-Threatening = Grade 4, Death = Grade 5)

- At least one Grade ≥ 3 Treatment-Related TEAE;
- At least one Serious TEAE;
- At least one Serious Treatment-Related TEAE;
- At least one TEAE leading to Discontinuation from Study Medication;
- At least one TEAE of Special Interest;
- At least one Treatment-Related TEAE of Special Interest

Except for all TEAEs which will be summarized by SOC and PT, the incidence of all other TEAEs by PT will be presented for the following

- All TEAEs by SOC and PT;
  - Treatment-related TEAEs (i.e., possibly or probably related);
  - All TEAEs by PT, and maximum severity (where the maximum intensity per subject will be counted at each level of summarization);
  - TEAES with Grade ≥ 3;

- Treatment-related TEAES with Grade ≥ 3;
- Serious TEAEs;
- Treatment-related serious TEAEs;
- Fatal TEAE;
- TEAEs leading to discontinuation from study treatment;
- TEAE of Special Interest;
- Treatment-related TEAE of Special Interest;
- Non-fatal serious TEAE;
- Hypertension TEAEs by Pre-existing Hypertension (Yes vs No) and Baseline Hypertension (Yes vs. No). Hypertension TEAEs will be selected using the same list of preferred terms as specified in Section 5.2.
- All TEAEs by PT occurring in at least 2% of subjects in the vibegron arm and greater than the placebo arm will be created and sorted by descending frequency in the vibegron arm.
- Gastrointestinal/IBS-related TEAEs by PT will be prepared and sorted by descending frequency in the vibegron arm. The final list of GI/IBS-related PT will be finalized and documented prior to database lock.

Adverse events of special interest for this study include:

- Adverse events suggestive of cystitis or urinary tract infection (UTI)
- Potential major cardiac and cerebrovascular events, including all cause death, myocardial infarction, cerebrovascular accident, hospitalization for unstable angina or chest pain, hospitalization for heart failure and coronary revascularization/angioplasty/stent
- Hypertension
- AEs consistent with orthostatic hypotension as confirmed by orthostatic vital signs
- Elevated serum aspartate aminotransferase (AST) or alanine aminotransferase (ALT)

lab value requiring that study drug be temporarily withheld or permanently discontinued

Overall AE listing will include the treatment arm, start and stop dates/times of the AE, first dose and last dose dates/times of double-blind treatment and days on study relative to the day of first dose and last dose of double-blind study treatment. In addition, the listing will indicate if the AE is treatment emergent (Yes vs No), an AE is of Special Interest (Yes vs No) and include the onset period (See Section 4.2.4.1 for the definition of the periods).

A Treatment related AE is defined as an AE for which the investigator classifies the AE as being "Probably Related" or "Possibly Related" to study treatment on Adverse Event CRF. Missing relationship and severity (intensity) will be imputed per Section 4.2.2.

The following additional listings will be provided:

- Listing of deaths
- Listing of Serious TEAEs
- Listing of treatment-emergent AESIs
- Listing of TEAEs leading to withdrawal or Interruption of study treatment

7.4. Laboratory Evaluations

The Clinical Laboratory Tests are the following:

Hematology	**Chemistry**	**Urine Dipstick/ Urinalysis[a]**	**Other**
Hematocrit Hemoglobin Platelet count WBC (total and differential) RBC	Albumin Alkaline phosphatase ALT AST Bicarbonate Calcium Chloride Creatinine[e] FSH Glucose Lipase Potassium Sodium Total bilirubin Direct bilirubin[f] Blood urea nitrogen Total cholesterol	Blood Glucose Protein Specific gravity Microscopic exam pH Color Urine pregnancy test (B-hCG)b	Serum β-hCG[b] Urine Drug screen Fecal calprotectin[c,d] Serum tissue transglutaminase antibody (IgA)[d] Coagulation (INR/PT/APTT)[g]

APTT = activated partial thromboplastin time; ALT = alanine aminotransferase; AST = aspartate aminotransferase;
β-hCG = β−human chorionic gonadotropin; INR = international normalized ratio; PT = prothrombin time; RBC = red blood cell count; WBC = white blood cell count
[a] A sample for urinalysis and urine culture/sensitivity testing will be sent to the laboratory only if the urine dipstick performed at the site tests positive for the presence of leukocytes, nitrites, or blood cells.
[b] Urine β-hCG will be tested for women of childbearing potential only. If urine β-hCG is positive, a serum β-hCG must be performed.
[c] Optional at Screening only.
[d] Performed at Screening only per inclusion criterion #8.
[e] Estimated glomerular filtration rate will be calculated and reported by the central lab.
[f] If total bilirubin is elevated above the upper limit of normal.
[g] Only upon request from Principal Investigator (if ALT, AST and bilirubin are increased).

All continuous laboratory parameters will be summarized descriptively by absolute value at each visit by treatment group, together with the corresponding changes from baseline. All parameters will be summarized in SI units. For continuous laboratory parameters with values that have a qualifier of "<" , " >", "≥", or "≤" than a limit of quantification will be set to the quantification value (e.g. <0.03 will be set to 0.03).

The number and percentage of subjects with laboratory measurements outside of the central laboratory normal range will also be summarized by treatment group and visit. Shift tables from baseline to maximum post-baseline value, to minimum post-baseline value, last post-baseline value, and at each post-baseline visit will be provided for hematology and chemistry parameters to display low, normal, high, and missing values by treatment group in a 3-by-3 contingency table. Denominators for percentages will be the number of subjects with non-missing data at the specific assessment and baseline.

Maximum post-baseline total bilirubin will be presented (<2 and ≥2 x ULN) and plotted against maximum post-baseline ALT (<3, ≥3 - <5, ≥5-<10, and ≥10 x ULN), expressed as multiples of ULN. This will be repeated to show maximum post-baseline total bilirubin against maximum post-baseline AST.

Data for subjects, who may potentially meet Hy's law, with ALT or AST ≥3 x ULN, and bilirubin ≥2 x ULN will be presented, which will include all visits for this subset of subjects. A line plot of liver biochemistry test results (including ALP, ALT, AST, total bilirubin, and GGT) over time will also be presented for this subset of subjects.

Urinalysis was performed at the central laboratory and the urine cultures were analyzed by local laboratories. A sample for the urinalysis and urine culture were assessed only if the urine dipstick performed at the site tests positive for the presence of leukocytes, nitrites, or blood cells. Urinalysis data will be summarized for each visit for both categorical and numeric results.

Any data outside the central laboratory normal reference ranges will be explicitly noted on the listings that are produced.

7.5. Vital Signs

Vital sign data including blood pressure, pulse rate and body temperature will be collected at all study visits except for the safety telephone follow-up visit. Body weight will be measured at Screening, baseline visit and Week 12 and Height will be measured at Screening only. Blood pressure and pulse will be measured in triplicate at each visit. The average of triplicates will be used for summary.

For all parameters, absolute values and change from baseline will be presented for scheduled visits using descriptive statistics for continuous variables. In addition, the maximum post-baseline, change from baseline to the maximum post-baseline, and change from baseline to the end of treatment will be summarized for each of the vital signs.

To further investigate changes in SBP/DBP/pulse rate from baseline, the following tables will be prepared:

- Counts and percent of subjects in each group with at least a 5/10/15 CFB at 3 consecutive post-baseline visits and at Week 12 will be produced for all subjects, by pre-existing hypertension category (Yes/No) and by baseline hypertension category (Yes/No). If the percentage of subjects in a subgroup is less than 25% of the total population, the subgroup in question will be not included in summary

- CFB in Maximum post-baseline (including 95% CI of Mean)

A by-subject listing, sorted by subject identifier, will be presented including all vital sign results (scheduled or unscheduled).

7.6. ECG

12-Lead ECG data will be collected at the screening visit only. All data collected will be listed.

7.7. Physical Examination

Brief physical examination data will be collected at Screening, Baseline, Week 12. Shift tables (normal, abnormal, not done) of baseline versus the last observation post-baseline [normal, abnormal (same as baseline), abnormal (new or aggravated), not done] may be generated, presenting the assessment for each component of the physical examination separately. Listing of abnormal results will be produced.

8. COVID-19 CONSIDERATIONS

This study was conducted during the COVID-19 global pandemic. During the time period of the pandemic, it is anticipated that changes in study visit schedules, missed visits, or subject discontinuations may lead to missing information (e.g., for protocol-specified procedures). In accordance with the FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Public Health Emergency, it is important to capture *specific* information in the case report form that explains the basis of the missing data, including the relationship to COVID-19 for missing protocol-specified information (e.g., from missed study visits or study discontinuations due to COVID-19).

The proportion of subjects with COVID-19 impact on visits will be summarized by treatment group overall and by visit for the FAS, FAS-D and FAS-M populations. The type of impact (e.g., visit cancelled, visit postponed, etc.) and the reason (e.g. subject acquired COVID-19, Subject unable to travel due to COVID-19, investigative site closure due to COVID-19, etc.) will also be summarized. A listing based on all subjects who had COVID-19 information collected will be generated which will document the visit impacted, type of impact, reason, date of contact of any telephone visits (if applicable), and if any doses were missed due to IP

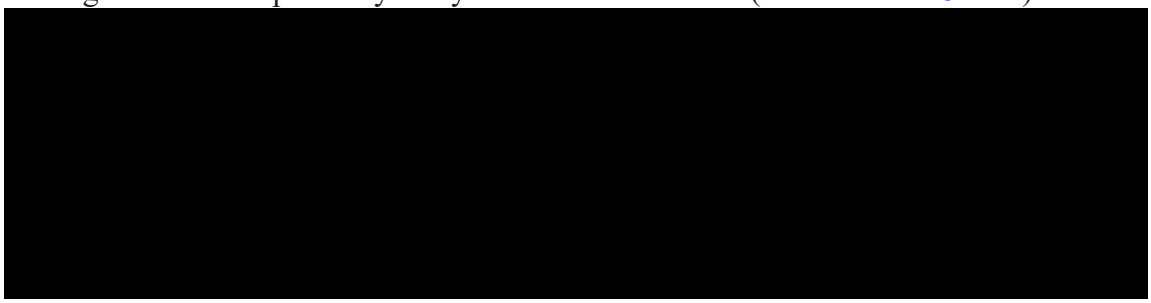
unavailability. Protocol deviations associated with COVID-19, identified with a "#COVID" in the protocol description, will also be included in the listing.

9. CHANGE FROM ANALYSIS PLANNED IN PROTOCOL

In Section 1.1.3, the other efficacy endpoint of the "Change from baseline to Weeks 2, 4, 8, and 12 in average weekly number of days of abdominal pain within 1 hour of eating for all IBS subjects including IBS-D and IBS-M subjects" was changed to "Change from baseline to weeks 2,4,8, and 12 in pain within 1 hour of eating for all IBS subjects including IBS-D and IBS-M subjects" since the data was collected as Yes or No to the following question, "In the past 7 days, did you experience abdominal pain within 1 hour of eating". The data will be summarized by visit as the proportion of subjects who improved vs not improved (no change, worsened, or missing post-baseline assessment).

Similarly, in Section 1.1.3, the "Change from baseline to weeks 2, 4, 8 and 12 in the weekly number of days with pain associated with a bowel movement" was changed to "Change from baseline to weeks 2, 4, 8 and 12 in pain associated with a bowel movement".

The following additional exploratory analyses will be conducted (See Section 4.3.1.11):



10. REFERENCES

[1] Kelleher DL, Hicks KJ, Cox DS, Williamson RR, Alpers DH, and Dukes GE. Randomized double-blind, placebo (PLA)-controlled, crossover study to evaluate efficacy and safety of the beta 3-adrenergic receptor agonist solabegron in subjects with irritable bowel syndrome. Neurogastroenterol Motil. 2008;20(Suppl 1):131.

[2] D. L. Patrick, D. A. Drossman, I. O. Frederick, J. DiCesare, and K. L. Puder, "Quality of life in persons with irritable bowel syndrome: development and validation of a new measure," *Dig. Dis. Sci.*, vol. 43, no. 2, pp. 400–411, Feb. 1998.

[3] "Information Sheet on the Irritable Bowel Syndrome-Quality of Life Measure (IBS-QOL)." [Online]. Available: http://depts.washington.edu/seaqol/docs/IBS-QOL_Info.pdf.

[4] "WPAI:SHP V2.0." [Online]. Available: http://www.reillyassociates.net/WPAI_SHP.html.

[5] "WPAI Scoring." [Online]. Available: http://www.reillyassociates.net/WPAI_Scoring.html.

[6] S. Greenland and J. M. Robins, "Estimation of a Common Effect Parameter from Sparse Follow-Up Data," *Biometrics*, vol. 41, no. 1, p. 55, Mar. 1985.

[7] D. B. Rubin, *Multiple imputation for nonresponse in surveys*. New York: Wiley, 1987.

[8] "Guideline on adjustment for baseline covariates in clinical trials." EMA, 26-Feb-2015.
[9] "Guidance for Industry for Irritable Bowel Syndrome - Clinical Evaluation of Drugs for Treatment." FDA, May-2012
[10] Patrick DL, Drossman DA, Frederick IO, et al. Quality of life in persons with irritable bowel syndrome: Development of a new measure. Dig Dis Sci 1998;43:400–11
[11] Chunlei Ke, Jianming Wang, Charlie Zhang, Qi Jiang & Steven Snapinn (2017): On Errors in Stratified Randomization, Statistics in Biopharmaceutical Research, DOI:10.1080/19466315.2016.1270229

11. APPENDIX

11.1. Table of Contents for Data Display Specifications

11.1.1. Output Tables

Table 8: List of Output Tables

Title		Analysis Set	Programming Note	Deliverable
	Study Population			
14.1.1.1	Subject Enrollment -All Subjects	Screened	Include screening and run-in. The percentages within the run-in should be based on the number of subjects who entered the run-in.	EOS
14.1.1.2.1	Subjects by Population and Subject Disposition	Randomized		TL, EOS
14.1.1.2.2-14.1.1.2.4	Subjects by Population and Subject Disposition	FAS-D/FAS-M/FAS		TL. EOS
14.1.1.3	Subject Randomization by US Investigator	Randomized		EOS
14.1.2.1.1-14.1.2.1.3	Major Protocol Deviations	FAS-D/FAS-M/FAS		EOS
14.1.2.2.1-14.1.2.2.3	Reasons for Exclusion from Analysis	FAS-D/FAS-M/FAS		EOS
14.1.3.1	Demographics and Baseline Characteristics	FAS	Race presented in descending frequency	TL, EOS
14.1.3.2	Demographics and Baseline Characteristics – IBS-D Subjects	FAS-D	Race presented in descending frequency	TL, EOS
14.1.3.3	Demographics and Baseline Characteristics – IBS-M Subjects	FAS-M	Race presented in descending frequency	TL, EOS

Title		Analysis Set	Programming Note	Deliverable
14.1.3.4	Demographics and Baseline Characteristics – IBS-D Subjects	PPS-D	Race presented in descending frequency	EOS
14.1.3.5	IBS Baseline Characteristics	FAS		TL, EOS
14.1.3.6	IBS Baseline Characteristics – IBS-D Subjects	FAS-D		TL, EOS
14.1.3.7	IBS Baseline Characteristics – IBS-M Subjects	FAS-M		TL, EOS
14.1.3.8	Demographics and Baseline Characteristics	Randomized		EOS
14.1.4.1	Medical History	FAS		EOS
14.1.4.2	Medical History	SAF		EOS
14.1.5.1	Prior Non-IBS Medication	FAS	ATC Level 2 and 4	EOS
14.1.5.2	Prior IBS Medication	FAS	ATC Levels 2, 4 and PT	EOS
14.1.5.3	Non-Rescue Concomitant Medication	FAS	ATC Level 2 and 4	EOS
14.1.5.4	Rescue Concomitant Medication	FAS	ATC Levels 2, 4 and PT	EOS
14.1.5.5	Additional Rescue Medication Data in Double-Blinded Treatment Period	FAS	Additional questions in CM form under ATC Levels 2	EOS
14.1.5.6	Prior IBS Medication	SAF		EOS
14.1.5.7	Non-Rescue Concomitant Medication	SAF		EOS
14.1.6.1-14.1.6.3	COVID-19 Impact	FAS-D/FAS-M/FAS		EOS
Efficacy Endpoints				
14.2.1.1	API Weekly Responder Primary Analysis (CMH): ≥30% Improvement) – IBS-D Subjects	FAS-D	Include all intervals but over Weeks 1-12 in the front	TL, EOS
14.2.1.2	API Weekly Responder Primary Analysis (CMH): ≥30% Improvement) – IBS-D Subjects	PPS-D	Include all intervals but over Weeks 1-12 in the front	EOS

Title		Analysis Set	Programming Note	Deliverable
14.2.1.3.1	API Weekly Responder over Weeks 1-12: ≥30% Improvement Sensitivity Analysis (Logistic Regression) – IBS-D Subjects	FAS-D		EOS
14.2.1.3.2	[redacted]	FAS-D	Include all intervals but over Weeks 1-12 in the front; The subject is considered a non-responder for the days receiving pain medication	EOS
14.2.1.4	API Weekly Responder Analysis (CMH): ≥40% Improvement–IBS-D Subjects	FAS-D	Include all intervals but over Weeks 1-12 in the front	TL, EOS
14.2.1.5	[redacted]	FAS-D	Include all intervals but over Weeks 1-12 in the front; The subject is considered a non-responder for the days receiving pain medication	EOS
14.2.1.6	API Weekly Responder Analysis (CMH): ≥50% Improvement–IBS-D Subjects	FAS-D	Include all intervals but over Weeks 1-12 in the front	TL, EOS
14.2.1.7	[redacted]	FAS-D	Include all intervals but over Weeks 1-12 in the front; The subject is considered a non-responder for the days receiving pain medication	EOS
14.2.1.8	API Weekly Responder Analysis (CMH): ≥ 30% Improvement -IBS-M Subjects	FAS-M		TL, EOS
14.2.1.9 – 14.2.1.10	[redacted]	FAS-M		EOS
14.2.1.11	API Weekly Responder Analysis (CMH): ≥ 30% Improvement -IBS All Subjects	FAS		TL, EOS
14.2.1.12 – 14.2.1.13	API Weekly Responder Analysis (CMH): ≥ 40%/50% Improvement -IBS All Subjects	FAS		EOS

Title		Analysis Set	Programming Note	Deliverable
14.2.1.14 – 14.2.1.16	API Weekly Responder over Weeks 1-12 Subgroup Analysis by API Strata (CMH): ≥ 30%/40%/50% Improvement – IBS-D Subjects	FAS-D	Percentages are based on the n within each subgroup	EOS
14.2.1.17	API Weekly Responder ≥ 30%/40%/50% Improvement over Weeks 1-12 by Subgroup	FAS	Descriptive summary. Subgroups are listed under SAP subgroup section; include three thresholds	EOS
14.2.2.1	Global Improvement Scale Responder Analysis (CMH) – IBS-D Subjects	FAS-D	Include all Weeks but Week 12 in the front. Include a footnote, "A responder is defined as a subject who answered that their symptoms were either moderately relieved or significantly relieved."	TL, EOS
14.2.2.3	Global Improvement Scale Responder Analysis (CMH) – IBS-M Subjects	FAS-M	Include all Weeks but Week 12 in the front	EOS
14.2.2.4	Global Improvement Scale Responder Analysis (CMH) –All Subjects	FAS	Include all Weeks but Week 12 in the front	EOS
14.2.2.5	Shift Table of Global Improvement Scale from Baseline - IBS-D Subjects	FAS-D	Include all Weeks but Week 12 in the front	EOS
14.2.2.6	Shift Table of Global Improvement Scale from Baseline – IBS-M Subjects	FAS-M	Include all Weeks but Week 12 in the front	EOS
14.2.2.7	Shift Table of Global Improvement Scale from Baseline – IBS-M Subjects	FAS	Include all Weeks but Week 12 in the front	EOS
14.2.3.1 – 14.2.3.3	Change from Baseline in Weekly Average Worst Abdominal Pain Score (MMRM) – IBS-D/IBS-M/All Subjects	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	TL, EOS
14.2.3.4	Change from Baseline in Weekly Average Worst Abdominal Pain Score	FAS	Place IBS-D, IBS-M and All IBS on the column and treatment arm underneath IBS columns	EOS
14.2.4.1 – 14.2.4.3	Change from Baseline in Weekly Average of Number of Pain Free Day (MMRM) – IBS-D/IBS-M/All Subjects	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS

Title		Analysis Set	Programming Note	Deliverable
14.2.4.4	Change from Baseline in Weekly Average of Number of Pain Free Day		Table 14.2.3.4	EOS
14.2.5.1 – 14.2.5.3	Change from Baseline in Abdominal Pain Within 1 Hour of Eating – IBS-D/IBS-M/All Subjects	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.5.4	Change from Baseline in Abdominal Pain Within 1 Hour of Eating	FAS	Table 14.2.3.4 but showing proportion of subjects Improved vs Not Improved (No Change, Worsened, Missing Post-baseline Assessment)	EOS
14.2.6.1 – 14.2.6.3	Change from Baseline in Abdominal Pain Associated with a Bowel Movement – IBS-D/IBS-M/All Subjects	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.6.4	Change from Baseline in Abdominal Pain Associated with a Bowel Movement	FAS	Table 14.2.3.4 but showing proportion of subjects Improved vs Not Improved (No Change, Worsened, Missing Post-baseline Assessment)	EOS
14.2.7.1 – 14.2.7.3	[redacted]	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.7.4		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.8.1 – 14.2.8.3		FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.8.4		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.9.1 – 14.2.9.3		FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.9.4		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS

	Title	Analysis Set	Programming Note	Deliverable
14.2.10.1 – 14.2.10.3	[redacted]	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.10.4		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.11.1 – 14.2.11.3		FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.11.4		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.12.1		FAS-D	Include all intervals but Week 1-12 in the front	EOS
14.2.14.1		FAS-D	Include all Weeks but Week 12 in the front	EOS
14.2.14.2		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.15.1		FAS-D	Include all Weeks but Week 12 in the front	EOS
14.2.15.2		FAS	Table 14.2.3.4	EOS
14.2.16.1		FAS-D	Include all Weeks but Week 12 in the front	EOS
14.2.16.2		FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.17.1.1 – 14.2.17.3.1		FAS-D/FAS-M/FAS	Include all intervals, but over Weeks 1-12 in the front	EOS
14.2.17.1.2-14.2.17.3.2		FAS-D/FAS-M/FAS	Include all intervals, but over Weeks 1-12 in the front	EOS

Title	Analysis Set	Programming Note	Deliverable
14.2.18.1 – 14.2.18.3 [redacted]	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front Include all 4 domains	EOS
14.2.18.4	FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front	EOS
14.2.19.1 – 14.2.19.3	FAS-D/FAS-M/FAS	Include all Weeks but Week 12 in the front	EOS
14.2.19.4 – 14.2.19.6	FAS-D/FAS-M/FAS	Include 8 domains; Include all Weeks but Week 12 in the front	EOS
14.2.19.7 – 14.2.19 .9	FAS-D/FAS-M/FAS	Include Week 12 first	EOS
14.2.19.10	FAS	Table 14.2.3.4; Include all Weeks but Week 12 in the front immediately following Baseline	EOS
14.2.19.11	FAS	Similar to Table 14.2.3.4 but showing proportion of subjects who responded	EOS
14.2.20.1-14.2.20.3	FAS-D/FAS-M/FAS		TL, EOS
14.2.20.4-14.2.20.6	FAS-D/FAS-M/FAS		EOS
14.2.20.7-14.2.20.9	FAS-D/FAS-M/FAS		EOS

	Title	Analysis Set	Programming Note	Deliverable
14.2.21.1-14.2.21.3	[redacted]	FAS-D/FAS-M/FAS		TL, EOS
14.2.21.4-14.2.21.6		FAS-D/FAS-M/FAS		EOS
14.2.21.7-14.2.21.9		FAS-D/FAS-M/FAS		EOS
14.2.22.1-14.2.22.3		FAS-D/FAS-M/FAS		TL, EOS
14.2.22.4-14.2.22.6		FAS-D/FAS-M/FAS		EOS
14.2.22.7-14.2.22.9		FAS-D/FAS-M/FAS		EOS
14.2.23.1-14.2.23.3	Missing Diary Days	FAS-D/FAS-M/FAS		TL, EOS
Safety Endpoints				
14.3.1.1	Treatment Exposure	SAF		EOS
14.3.1.2	Treatment Compliance	SAF		EOS
14.3.1.3	Treatment Exposure	All Subjects	To capture the subject with a dosing error during run-in	
Adverse Events				

Title		Analysis Set	Programming Note	Deliverable
14.3.2.1	Overall Treatment-Emergent Adverse Events	SAF	Include the AESI categories as subcategories	TL, EOS
14.3.2.2	Treatment-Emergent Adverse Events by System Organ Class and Preferred Term	SAF		EOS
14.3.2.3	Treatment-related TEAEs by Preferred Term	SAF		EOS
14.3.2.4	Treatment-Emergent Adverse Events by Preferred Term and Maximum Intensity	SAF		EOS
14.3.2.5	Treatment-Emergent Adverse Events Occurring in ≥ 2% Subjects of Vibegron Arm and Greater Than the Placebo by Preferred Term	SAF		EOS
14.3.2.6	Treatment-Emergent Adverse Events with Grade ≥ 3 by Preferred Term	SAF		TL, EOS
14.3.2.7	Treatment-related TEAEs with Grade ≥ 3 by Preferred Term	SAF		EOS
14.3.2.8	Serious TEAEs by Preferred Term	SAF		EOS
14.3.2.9	Treatment-related Serious TEAEs by Preferred Term	SAF		EOS
14.3.2.10	Fatal TEAEs by Preferred Term	SAF		EOS
14.3.2.11	TEAEs Leading to Treatment Discontinuation by Preferred Term	SAF		TL, EOS
14.3.2.12	TEAEs of Special Interest by Preferred Term	SAF		EOS
14.3.2.13	Treatment-related TEAEs of Special Interest by Preferred Term	SAF		TL, EOS
14.3.2.14	Non-Fatal Serious TEAEs by Preferred Term	SAF		EOS

Title		Analysis Set	Programming Note	Deliverable
14.3.2.15	Hypertension TEAEs by Preferred Term and Pre-existing Hypertension	SAF	Search for the following terms hypertension PTs: Accelerated hypertension, Diastolic hypertension, Essential hypertension, Hypertension, Hypertensive crisis, Hypertensive emergency, Hypertensive heart disease, Malignant hypertension, Malignant hypertensive heart disease, Secondary hypertension, Supine hypertension, Systolic hypertension	EOS
14.3.2.16	Hypertension TEAEs by Preferred Term and Baseline Hypertension Status	SAF	Search for the following terms hypertension PTs: Accelerated hypertension, Diastolic hypertension, Essential hypertension, Hypertension, Hypertensive crisis, Hypertensive emergency, Hypertensive heart disease, Malignant hypertension, Malignant hypertensive heart disease, Secondary hypertension, Supine hypertension, Systolic hypertension	EOS
14.3.2.17	GI/IBS-related TEAEs by Preferred Term	SAF		EOS
14.3.2.18	List of Deaths	SAF		EOS
14.3.2.19	Listing of Treatment Emergent Serious TEAEs	SAF		EOS
14.3.2.20	Listing TEAEs Leading to Withdrawal or Interruption of Study Treatment	SAF		EOS
14.3.2.21	Listing of TEAEs of Special Interest	SAF		EOS

Title		Analysis Set	Programming Note	Deliverable
Laboratory data				
14.3.3.1	Hematology Laboratory Parameters	SAF	Include observed and CFB; include conventional units as well for hematocrit	EOS
14.3.3.2	Abnormal Classification of Hematology Laboratory Parameters	SAF		EOS
14.3.3.3	Shift Table of L/N/H Classification for Hematology Laboratory Parameters from Baseline	SAF	Include max/min post-dose, by visit, last-post dose	EOS
14.3.3.4	Clinical Chemistry Laboratory Parameters	SAF	Include observed and CFB	EOS
14.3.3.5	Abnormal Classification of Clinical Chemistry Laboratory Parameters	SAF		EOS
14.3.3.6	Shift Table of L/N/H Classification for Clinical Chemistry Laboratory Parameters from Baseline	SAF		EOS
14.3.3.7	Maximum Post-baseline ALT and AST vs. Maximum Post-baseline Bilirubin	SAF		EOS
14.3.3.8	Listing of Subjects Potentially Met Hy's Law	SAF		EOS
Other Safety				
14.3.4.1	Vital Sign Parameters	SAF	Include all observed and CFB	EOS
14.3.4.2	Vital Sign Parameter Change from Baseline Shifts at 3 Consecutive Visits	SAF	SBP, DBP and PR	EOS
14.3.4.3	Vital Sign Parameter Change from Baseline Shifts at Week 12	SAF		EOS
14.3.4.4	Vital Sign Parameter Maximum Post-Baseline Change from Baseline	SAF		EOS
14.3.5.1	Physical Examination Shift from Baseline	SAF		EOS

[1] TL= Topline Results (priority delivery after DBL); EOS= End of Study

11.1.2. Output Figures

Table 9: List of Output Figures

Title		Population	Programming Note	Deliverable
Study Population				
14.1.1	Kaplan Meier Plot of Time to Study Discontinuation	SAF		EOS
Efficacy Endpoints				
14.2.1.1	Forest Plot of Risk Difference Estimates and 90% CI API Weekly Responder Analysis over Weeks 1-12	FAS	Include all thresholds. Under each threshold, list IBS-D, IBS-M and All IBS in order	TL, EOS
14.2.1.2	Line Plot of Proportion Subjects Who Met Weekly Responder with ≥30% improvement by Week – IBS-D Subjects	FAS	Include all individual weeks.	EOS
14.2.1.3	Line Plot of Proportion Subjects Who Met Weekly Responder ≥40% improvement by Week – IBS-D Subjects	FAS	Include all individual weeks.	EOS
14.2.1.4	Line Plot of Proportion Subjects Who Met Weekly Responder ≥50% improvement by Week – IBS-D Subjects	FAS	Include all individual weeks.	EOS
14.2.1.5	Line Plots of LSMEAN (SE) of Change from Baseline in Weekly Average Worst Abdominal Pain Score from MMRM – IBS-D Subjects	FAS	Include Weeks 2, 4, 8 and 12	EOS
14.2.1.6	Line Plots of LSMEAN (SE) of Change from Baseline in Weekly Average Worst Abdominal Pain Score from MMRM – IBS-M Subjects	FAS-M	Include Weeks 2, 4, 8 and 12	EOS
14.2.1.7	Line Plots of LSMEAN (SE) of Change from Baseline in Weekly Average Worst Abdominal Pain Score from MMRM – FAS Subjects	FAS	Include Weeks 2, 4, 8 and 12	EOS

	Title	Population	Programming Note	Deliverable
14.2.1.8	[redacted]	FAS	Include all thresholds. Under each threshold, list IBS-D, IBS-M and All IBS in order	EOS
14.2.1.9	Forest Plot of Risk Difference Estimates and 90% CI API Weekly Responder AND < 50% Increase from Baseline in Average Number of Days/Week with BSS Type 6 or 7 over Weeks 1-12	FAS	Include all thresholds. Under each threshold, list IBS-D, IBS-M and All IBS in order	EOS
14.2.1.10	[redacted]	FAS	Include all thresholds. Under each threshold, list IBS-D, IBS-M and All IBS in order	EOS
Safety Endpoints				
14.3.3.1	Scatter Plot of Maximum Post-Baseline ALT versus Maximum Total Bilirubin Expressed as Multiples of ULN	SAF		EOS
14.3.3.2	Scatter Plot of Maximum Post-Baseline AST versus Maximum Total Bilirubin Expressed as Multiples of ULN	SAF		EOS
14.3.3.3	Line Plot of Liver Chemistry Test Results over Time for Subjects with Elevated ALT or AST, and Elevated Total Bilirubin at Any Time	SAF	Only produce for subjects who potentially met Hy's law	EOS

11.1.3. Output Listings

Table 8: List of Output Listings

Title		Population	Programming Note	Deliverable
Disposition and Demographics				
16.2.1.1	Subject Disposition for Screen Failures	Screened		EOS

Title		Population	Programming Note	Deliverable
16.2.1.2	Run-In Treatment Administration	Placebo Run-In Set		EOS
16.2.1.3	Subject Disposition	Randomized		EOS
16.2.1.4	Subjects who did not Satisfy Inclusion/Exclusion Criteria	Screened		EOS
16.2.1.5	Subject Randomization Details	Randomized		EOS
16.2.1.6	Major Protocol Deviations	Randomized		EOS
16.2.1.7	Subject Demographic and Baseline Characteristics	Randomized		EOS
16.2.1.8	Medical History	Randomized		EOS
16.2.1.9	Prior and Concomitant Medications	Randomized		EOS
16.2.1.10	Prior Irritable Bowel Syndrome (IBS) Medications	Randomized		EOS
16.2.1.11	Additional Rescue Concomitant Medication information	Randomized		EOS
16.2.1.12	Subjects with Misclassification of Stratification Factors	Randomized		EOS
16.2.1.13	COVID-19 Impact	All Subject	include treatment IBS subtype and all the information collected on COVID-19 eCRF	EOS
Efficacy				EOS
16.2.2.1	Weekly API Responder	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata, Week	TL
16.2.2.2	Global Improvement Scale Responder Classification	FAS		EOS
16.2.2.3	Derived Parameters from Bowel Movement Diary	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata	EOS
16.2.2.4	Additional Abdominal Pain Questionnaires	FAS		EOS
16.2.2.5	Additional IBS-related Baseline Data			
16.2.2.6	[redacted]	FAS		EOS

Title		Population	Programming Note	Deliverable
16.2.2.7	[redacted]	FAS		EOS
16.2.2.8	Missing Diary Days	FAS		TL
16.2.2.9	Cumulative Weekly API Responder	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata	TL
16.2.2.10	[redacted]	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata	TL
16.2.2.11	[redacted]	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata	TL
16.2.2.12	Cumulative Daily API Responder	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata	TL
16.2.2.13	[redacted]	FAS	Sort by Treatment, IBS Subtype, Subject ID, Baseline API Strata	TL
Safety				
16.2.3.1.1	Treatment Exposure and Compliance	SAF		EOS
16.2.3.1.2	Subjects with Compliance > 120% who Were Not Lost to Follow-Up	SAF	To identify any subject with >> 120% compliance who was not lost of follow-up	EOS
16.2.3.2.1	Adverse Events	SAF		EOS
16.2.3.2.2	TEAE Serious Adverse Events	SAF	Use the vertical format	EOS
16.2.3.2.3	TEAEs of Special Interest	SAF	Use the vertical format	
16.2.3.3	Hematology Laboratory Results	SAF		EOS
16.2.3.4	Clinical Chemistry Laboratory Results	SAF		EOS
16.2.3.5	Urinalysis Laboratory Results	SAF		EOS

Title		Population	Programming Note	Deliverable
16.2.3.6	Vital Sign Results	SAF		EOS
16.2.3.7	ECG Results	SAF		EOS
16.2.3.8	Physical Examination	SAF		EOS
16.2.3.9	Other Laboratory Results	SAF		EOS

11.2. General Information Related to Data Presentation

The general formats and layouts of tables, listings, and figures (TLFs) are provided in a separate "programming consideration" document. Actual formats and layouts may be altered slightly to accommodate actual data or statistics. Minor format changes will not require updates to the statistical analysis plan (SAP).

The TLF numbering and general content follow the ICH E3 guidelines.

11.3. Summary of SAP Changes from v1.0 to v2.0

	Section	Change
#1	3.1.5	Modified definition of FAS to be at least one evaluable post-baseline weekly API score (i.e., where evaluable is considered minimum of 4 diary entries in a week).
	3.1.6	Removed an example from the paragraph regarding PPS
	3.2.1	Clarified that "significant" and "major" protocol deviations are the synonymous.

	Section	Change
#2	4.1.3	Added a subgroup of Baseline Bristol Stool Scale (< 5.5 vs ≥ 5.5)
	4.3.18	[redacted]
	4.3.4	Updated baseline definition to be the last recorded value on or prior to the date of double-blind treatment start (note date of randomization)
#3	5.4 5.5	Changed the medDRA version to "23.0" or higher Added "or later" after "World Health Organization (WHO) Drug B3-March Format, 2020 version" in case the dictionary version is changed before data base lock
#4	7.4	Clarified to set the value to the limit of quantification only for continuous laboratory parameters with values that have a qualifier of "<" , " >", "≥", or "≤".
#5	8	Added the protocol deviations associated with COVID to the listing

	Section	Change
#6	11.1	Updated the topline tables Added a shift table for GIS Added the daily composite responder analysis for IBS-M and IBS, respectively Removed the Summary of Urinalysis. The data will be listed since it is only available for subjects who had a positive dipstick
		[redacted]
		Other minor update