A Phase 2a, Randomized, Double-blind, Placebo-controlled, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of BOS-589 in the Treatment of Patients with Diarrhea-predominant Irritable Bowel Syndrome (IBS-D)

ClinicalTrials.gov Identifier: NCT03977155

Statistical Analysis Plan, Version 1.2, dated 12-November-2020
Statistical Analysis Plan, Version 1.1, dated 27-March-2020
Statistical Analysis Plan, Version 1.0, dated 29-January-2020

A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE EFFICACY, SAFETY, AND TOLERABILITY OF BOS-589 IN THE TREATMENT OF PATIENTS WITH DIARRHEA-PREDOMINANT IRRITABLE BOWEL SYNDROME (IBS-D)

Statistical Analysis Plan

VERSION 1.2
DATE OF PLAN:

Nov 12, 2020

STUDY DRUG:
BOS-589

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Approval Signature Page: PPD Inc.

PPD

PPD

PPD

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD

11/15/2020

Date

PPD

11/15/2020

Date

PPD

Modification History

After approval of Version 1.0 of the Statistical Analysis Plan, subsequent versions should be documented below with a brief description of the change from the previous version as well as the rationale for the change.

Version, Date	Made by	Brief Description of Change and Rationale
Version 1.1, 27-MAR-2020	PPD	Baseline definition in section 7 was revised to be based on the 7 days prior to the date the noted eligibility criteria were met.
Version 1.2, 11-NOV-2020	PPD	Definition of data used to calculate the average of daily diary entries to be used for analysis at each visit revised to be the week prior to the actual visit date for patients. Clarification made for references to the timing of collection for IBS-SS and IBS-GS.

Contents

ABBREVIATIONS

°C	Degrees Celsius
AE	Adverse event
ATC	Anatomic Therapeutic Chemical
AUC	Area under the curve
AUC_{0-4}	Area under the concentration versus time curve from time zero to 4 hours postdose
AUC_{0-t}	Area under the concentration versus time curve from time zero to the last quantifiable concentration
bid	Twice daily
BMI	Body mass index
BSFS	Bristol Stool Form Score
cm	Centimeters
C_{max}	Maximum observed plasma concentration
C_{min}	Minimum observed plasma concentration
CTCAE	Common Terminology Criteria for Adverse Events
ECG	Electrocardiogram
eCRF	Electronic case report form
IBS	Irritable bowel syndrome
IBS-GS	Irritable Bowel Syndrome Global Scale
IBS-D	Irritable bowel syndrome-diarrhea
IBS SS	Irritable Bowel Syndrome Severity Score
ITT	Intention-To-Treat Analysis Set
IWRS	Interactive web response system
kg	Kilograms
m	Meters
MedDRA	Medical Dictionary for Medical Affairs
mg	Milligrams
mm	Millimeters
mmHg	Millimeters of mercury
msec	Milliseconds
NTEAE	Non-treatment emergent adverse event
PD	Progressive disease or Pharmacodynamics
PK	Pharmacokinetics
PPS	Per- Protocol Analysis Set
PT	Preferred Term
Q1	25^{th} Percentile/First quartile
Q3	75^{th} Percentile/Third quartile
SAE	Serious adverse events
SAF	Safety
SAP	Statistical analysis plan
SAS®	Statistical Analysis System
SD	Standard deviation
SI	International System of Units

SMC	Safety monitoring committee
SOC	System Organ Class
TEAE	Treatment-emergent adverse events
T_{max}	Time to maximum observed plasma concentration
WAP	Worst abdominal pain
WHO	World Health Organization

1. INTRODUCTION

The statistical analysis plan (SAP) details the planned analysis required to satisfy the Clinical Study Report of study number BOS-589-201: "A Phase 2a, Randomized, Double-blind, Placebo-controlled, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of BOS-589 in the Treatment of Patients with Diarrhea-predominant Irritable Bowel Syndrome (IBS-D)". The content of this SAP is based on the protocol amendment 2 (version 3.0) dated 12 Aug 2019.

Mock shells for tables, listings, and figures are in a separate document. Analyses for pharmacokinetics (PK) and pharmacodynamics (PD) will be governed by a separate SAP.

2. STUDY OBJECTIVES AND ENDPOINTS

OBJECTIVES	ENDPOINTS
Primary	
• To evaluate in patients with IBS-D the abdominal pain response to BOS-589 after 4 weeks of treatment, relative to Placebo.	• 24-hour worst abdominal pain scores (WAP) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the overall safety and tolerability of BOS-589 in the treatment of IBS-D during 4 weeks of treatment, relative to Placebo.	• Incidence of adverse events (AEs) serious adverse events (SAEs), discontinuations because of AEs, and any treatment-related severe AEs.
Secondary	
• To evaluate the treatment effect of BOS-589 on defecation after 4 weeks, relative to Placebo.	• Change in stool consistency, measured by the daily Bristol Stool Form Score (BSFS) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). • Change in stool frequency, measured by the total number of spontaneous bowel movements in 24 hours at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the treatment effect of BOS-589 on IBS-related signs and symptoms.	• Change in the IBS Severity Score (IBS-SS) at Day 29 compared to baseline. • Change in the IBS Global Scale (IBS-GS) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the steady-state PK of BOS-589.	• Maximum observed plasma concentration (C_{max}); time to C_{max} (T_{max}); minimum plasma concentration (C_{min}); area under the concentration versus time curve (AUC) from time zero to 4 hours postdose (AUC_{0-4}); AUC from time zero to the last quantifiable concentration (AUC_{0-t}).

Additional Objectives	Additional Endpoints
Exploratory (not included in protocol amendment 2 [version 3.0] dated 12 Aug 2019)	
• To explore the treatment effect of BOS-589 in terms of weekly responders after 4 weeks of treatment, relative to Placebo.	• Number of patients who are weekly responders, defined by both an abdominal pain intensity weekly responder (decrease in the weekly

Additional Objectives	Additional Endpoints
	average of the WAP in the past 24 hours score of at least 30%compared with baseline averaged over the week prior) and a stool consistency weekly responder (at least 50%reduction in the number of days per week with at least one stool that has a consistency of Type 6 or 7 compared with baseline averaged over the week prior), at each scheduled visit.

3. STUDY DESIGN

3.1. Study Design and Population

This study is a phase 2a, randomized, double-blind, Placebo-controlled, multicenter trial designed to provide the first proof-of-principle efficacy of BOS-589 in IBS-D patients, and to inform dose selection for subsequent development. The study will consist of a pre-treatment phase, a 4-week double-blind treatment phase, and a 2-week post-treatment follow-up period.

Pre-Treatment Phase: During the pre-treatment phase, patients will be evaluated for up to 5 weeks to assess eligibility. The pre-treatment phase will consist of initial screening assessments and a Run-in period.

After the initial screening assessments have been performed, eligible patients will enter a Run-in period of up to 3 weeks. During the Run-in period, the patients will complete an electronic diary to collect daily information related to their IBS-D symptoms, bowel function, and rescue medicine use, to confirm disease activity and diary compliance. Patients will also be requested to discontinue any prohibited medications during this phase of the study.

Upon completion of the Run-in period, patients will return to the study site to confirm eligibility for randomization into a 4-week double-blinded Treatment Phase.

Treatment Phase: Eligible patients will be randomized (1:1:1) into the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for 200 mg BOS-589 twice daily (bid) orally;
- Cohort 2 (Low Dose): 2 x 25 mg tablets for 50 mg BOS-589 bid orally;
- Cohort 3 (Placebo): 2 x visually matching Placebo oral tablets bid.

During the 4 weeks of double-blind treatment, patients will continue to record their daily IBS-D symptoms, bowel function, and rescue medicine use in the electronic diary, as described in Section 8.3 of the protocol.

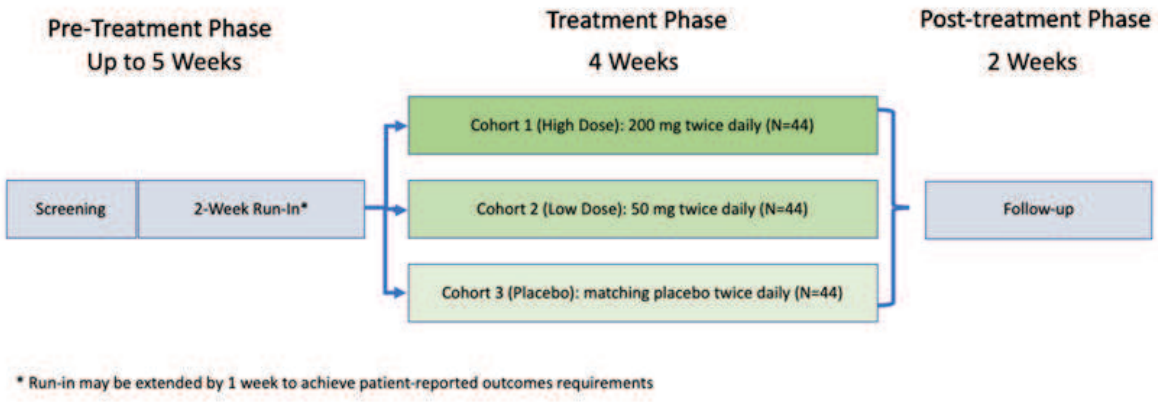
Post-treatment Phase: A 2-week post-treatment follow-up visit will occur for patients who complete the Treatment Phase. During the 2 weeks of follow-up, patients should continue to record their daily IBS-D symptoms, bowel function, and rescue medicine use in the electronic diary.

Patients who prematurely discontinue treatment should return to the study center to complete the early termination assessments as soon as possible after stopping the treatment.

Data analyses will occur after all patients in the trial have completed the last visit or procedure shown in the Schedule of Activities, Section 1.3 of the protocol.

Post-study Access to Therapy: No post-treatment access to therapy will be provided to patients randomized in the study.

Figure 1. Study Schematic



Further information on the study design, including eligibility criteria and a schedule of assessments, are detailed in the study protocol.

3.2. Changes to the Analyses from the Protocol

One exploratory efficacy endpoint was added in addition to the endpoints defined in the protocol.

- Summaries of weekly responders (defined in section 7.3) will be summarized as an exploratory efficacy endpoint.

The wording of how randomization file is handled is clarified in the SAP by saying the file will be held by an independent unblinded Biostatistician, not the study statistician.

3.3. Randomization and Blinding

This is a randomized, double-blinded, Placebo-controlled study. Patients should be randomized as close as possible to the time of the planned first dose of study drug.

Randomization will be conducted via an interactive web response system (IWRS) in a 1:1:1 ratio. Eligible patients will be randomized (1:1:1) into the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for 200 mg BOS-589 twice daily (bid) orally;
- Cohort 2 (Low Dose): 2 x 25 mg tablets for 50 mg BOS-589 bid orally;
- Cohort 3 (Placebo): 2 x visually matching Placebo oral tablets bid.

Patients will be assigned to the cohorts using a permuted block randomization stratified by site (44 patients per cohort). The randomization schedule will be generated by an independent unblinded Biostatistician and will be transferred to the IWRS team for loading into the system. PPD blinded biostatistics and blinded programming team will remain blinded. The randomization file will be held by the independent unblinded Biostatistician until the end of the trial and when the database has been locked.

The sponsor, site staff, and patients will be blinded to the treatment assignment. BOS-589 and Placebo will be similar in appearance, and provided in white, opaque, high-density polyethylene bottles, each with a child-resistant closure. Each bottle will be labeled as per country requirement.

Only in a medical emergency will a patient's treatment assignment be unblinded, and this process will be performed and documented in the IWRS. Every effort should be made to contact the Medical Monitor to discuss unblinding prior to breaking the blind.

3.4. Sample Size Considerations

This is a phase 2a study in which a hierarchical hypothesis test will be employed for the primary efficacy endpoint. The primary endpoint is defined as a change in the 24-hour WAP score at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). Mean change from baseline will be measured as a continuous variable and compared between cohorts in a hierarchical testing order using a t-test of equal variances and a two-tailed alpha of 0.05.

The sample size was based on the first hypothesis test in the hierarchical plan, with approximately 80% power, two-sided alpha of 0.05, and a standard deviation (SD) of 3 points on the numeric rating scale to detect a 1.6 minimum change between cohorts (N = 120 patients). To account for attrition and a potential higher Placebo rate, a total of 132 patients will be randomized to one of 3 cohorts in a 1:1:1 ratio (44 patients per cohort).

3.5. Safety Monitoring Committee

There is no safety monitoring committee reviewing data for this study.

3.6. Interim Analysis

No formal efficacy interim analysis is planned for the purposes of statistically comparing the data prior to the completion of the study.

3.7. Timing of Analyses

Formal hypothesis testing will occur at the end of the study and post database lock. There are no planned Data, Safety, or Adjudication Committees for this study and no planned statistical interim analyses.

4. DATA ANALYSIS CONSIDERATIONS

Unless noted otherwise, the following statistical considerations will be applied:

- All analyses will be conducted based on Statistical Analysis System (SAS®) software version 9.4 or higher.
- Data for all sites/centers will be pooled for analysis.
- All data in the database except birth date will be presented in by-patient data listings.
- Listings will be sorted by cohort, patient identification number (concatenated site and patient number), and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.
- Continuous data will be summarized by cohort (or last assigned dose level, as applicable) based on n, mean, median, standard deviation (SD), 25th percentile/first quartile (Q1), 75th percentile/third quartile (Q3), minimum value, and maximum value.
- Categorical data will be summarized by cohort using n and percentage based on the number of non-missing values.
    - The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
    - Counts of zero will be presented without percentages.
- Statistics will be presented in the summary tables based on the following:
    - Mean, Median, Q1, and Q3: one additional decimal place to that reported for Minimum and Maximum
    - SD: two additional decimal places than the Minimum and Maximum

    - Percentages: reported to one decimal place
    - P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.
- The number and percentage will be presented in the form of XX (XX%). Percentage values of less than 1% will be presented as XX (<1%).
- Statistical inference will be based on a 5% significance level (i.e., 95% confidence intervals [CIs] will be produced). CIs will be produced only in cases with sufficient sample size (i.e., n => 5).
- Change from baseline = value (post-baseline visit) – value (baseline or last visit prior to treatment).
- All data up to the time of study completion/withdrawal from the study will be included in the analysis, regardless of duration of treatment.
- Numbering for data displays will be based on the International Council for Harmonization E3.
- The protocol uses the terms cohort, treatment group, and treatment arm interchangeably. For consistency in this document, the term cohort will be used.

4.1. Stratification and Covariates

There are no formal plans for analysis stratification or use of covariates.

4.2. Evaluation of Subgroups

No subgroup analyses are powered for the study. However, analyses for the primary and secondary endpoints will be analyzed by sex, and fecal calprotectin < 50 mcg/g at screening, and possibly other subgroups based on emerging data.

4.3. Multiple Comparisons and Multiplicity

A hierarchical hypothesis test will be employed for the primary efficacy endpoint, in the following pre-defined testing order using a t-test of equal variances and a two-tailed alpha of 0.05:

Hypothesis 1:

Active Treatment (Cohort 1 + Cohort 2) versus Placebo (Cohort 3);

If statistically significant at P < 0.05, then proceed to Hypothesis 2.

Hypothesis 2:

Cohort 1 (High Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05, then proceed to Hypothesis 3.

Hypothesis 3:

Cohort 2 (Low Dose) versus Cohort 3 (Placebo);

If statistically significant at $P < 0.05$ then, proceed to Hypothesis 4.

Hypothesis 4:

Cohort 1 (High Dose) versus Cohort 2 (Low Dose).

P-values presented in the analyses of secondary endpoints are displayed for descriptive purposes only and are not adjusted for multiplicity.

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Assigned and Actual Treatment

Patients will be randomized via IWRS in a 1:1:1 ratio to receive BOS-589 200 mg, BOS-589 50 mg, or Placebo. Patients will be assigned to the cohorts using a permuted block randomization stratified by site (44 patients per group).

The assigned treatment is determined by a randomization schedule. The randomization schedule will be generated and will be transferred to the IWRS team for loading into the system. The actual treatment given to individual patient is the actual treatment.

All safety analyses will be based on the actual treatment. Study patient data and all efficacy analyses will be based on the assigned treatment unless specified otherwise.

5.2. Reference Dates

Reference dates are defined as follows:

- Screening date is defined as the electronic case report form (eCRF) provided visit date on which a patient was screened for trial entry.
- Informed consent date is defined as the eCRF provided date of informed consent from the informed consent page.
- Run-in period start date is defined as earliest date of eCRF provided visit date from the Run-in period page.
- Run-in period end date is defined as eCRF provided date of randomization -1 from the randomization page.
- Treatment start date is defined as the date of first dose of any study drug.
- Treatment end date is defined as the date of last dose of any study drug.
- Age will come directly from the eCRF which is integrated from IWRS based on the informed consent date.

- Safety data, such as AEs and laboratory assessments, will use the treatment start date as a reference date.
- Efficacy data will use the treatment start date as a reference date.
- Study day will be based on treatment start date as a reference date. Study Day 1 will be the day patients start the first dose of study drug.

5.3. Study Day and Duration Variables

Reference date calculations will generally be defined as the following, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest >= reference date;
- otherwise, date of interest – reference date.

If either date is missing, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.6.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – informed consent date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug. Patients still receiving ongoing treatment or participating in study follow up at the time of analysis will use imputed treatment end and end of study dates as described in Section 5.7.

Durations reported in days will be calculated as measure date – reference date + 1. When reporting in months the result will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.4. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date. Treatment end date is defined as the date of last dose of study drug.
- Post-treatment is defined as the period of time following the on-treatment period.

5.5. Baseline and Post-Baseline Changes

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value (including unscheduled assessments) collected prior to or on the treatment start date/time for measures collected prior to the treatment.
- Post-baseline values will be those collected after the date/time of first dose of study drug.
- Change from baseline is defined as: value – baseline value.
- Percentage change from baseline is defined as: (value – baseline value)/baseline value x 100%.
- Maximum change will be based on the change from baseline value to the maximum value for all post-baseline assessments collected during the on-treatment period, scheduled or unscheduled, using the formulas above.
- Minimum change will be based on the change from baseline value to the minimum value for all post-baseline assessments collected during the on-treatment period, scheduled or unscheduled, using the formulas above.

5.6. Imputation of Partial Dates

Appendix 1 details partial date conventions for AEs and medications.

Treatment End Date

Missing treatment end dates will not be imputed at the end of the study. However, due to ongoing reporting needs, treatment end date will be imputed as earliest of the data cutoff date, date of death, or last treatment date recorded on the exposure eCRF.

End of Study Date

Missing end of study dates will not necessarily be imputed at the end of the study. However, due to ongoing reporting needs, end of study dates will be imputed as the earliest of the data cutoff date, date of death, or last date recorded on the eCRF.

5.7. Multiple Assessments and Visit Windows

If there are multiple assessments at a scheduled visit, then the last assessment will be used in the analysis.

Nominal visits (e.g., those identified by the study eCRF) will be the basis of summarization and statistical analysis; visit windows for each scheduled visit can be found in the Schedule of Activities from the protocol. Unscheduled data may be included in summaries of baseline and maximum/minimum change values; summaries of specific abnormalities any time post-baseline; and patient data listings.

5.8. Lost to Follow Up or Lapse of Adequate Assessments

A patient will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site staff. Missing evaluations will not be imputed.

5.9. Cohort Display

Cohorts will be displayed with the following columns. See the mock shells for additional details and a visual representation of the cohort display.

- 200 mg BOS-589 bid
- 50 mg BOS-589 bid
- Active Treatment (Combined 200 mg and 50 mg Cohorts)
- Placebo bid
- Overall (not needed for efficacy outputs)

For primary efficacy analyses testing, cohort significant p-values will be displayed in the hierarchical order below. Following the first p-value, each subsequent p-value will only be displayed if the prior p-value is < 0.05.

- Active Treatment (Combined 200 mg and 50 mg Cohorts) vs. Placebo
- 200 mg BOS-589 bid vs. Placebo
- 50 mg BOS-589 bid vs. Placebo
- 200 mg BOS-589 bid vs. 50 mg BOS-589 bid

5.10. Missing Data

Missing data imputations for efficacy endpoints are described in Section 7; AE and concomitant medication, ongoing end of treatment, and ongoing end of study are described in Section 5.6. Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Sets

- The Screened Analysis Set will include all patients who have signed an informed consent document.
- The Intention-to-Treat (ITT) Analysis Set will include all randomized patients.
- The Per Protocol (PP) Analysis Set will include all patients in the ITT Analysis Set who received at least 1 dose of study drug and did not have any major protocol deviations.
- The Safety (SAF) Analysis Set will include all patients who received at least 1 dose of study drug.

All safety analyses will be performed in the SAF Analysis Set by the actual cohorts. All efficacy analyses will be performed in the ITT and PP Analysis Sets by the randomized cohorts. Summaries of the Screened Analysis Set will be performed by the randomized cohorts.

The criteria for inclusion in the PP Analysis Set will be finalized and documented before database is locked and study unblinding. Additional information on inclusion into this population can be found in Section 6.3.

6.2. Patient Disposition

A summary of patient enrollment by site will be presented by cohort and overall for the Screened Set. The overall patient enrollment status (screened [i.e., consented], screen failures, randomized) will be provided. Patients who sign the ICF and are randomized but do not receive the study drug may be replaced. Once randomized, patients who are withdrawn from therapy before completing 28 days of dosing or are discontinued from the study for any reason will not be replaced.

Disposition data will be summarized for the Screened Set. The number of patients in each analysis set (ITT, SAF, PPS), completed treatment, with ongoing treatment, completed the study, and still ongoing in the study, will be summarized. The number of patients who discontinued study drug including reasons, the duration on study drug, and duration on study will be summarized. The number of patients who discontinued the study including reasons, the duration on study drug, and duration on study will also be summarized. Data will be presented by cohort and overall by assigned treatment.

A by-patient listing of patient enrollment data as well as a by-patient listing of patient disposition data including reason for treatment and study discontinuation will be presented for the screened patients.

6.3. Protocol Deviations

Protocol deviations will be recorded in the clinical trial management system.

Protocol deviations will be identified and classified as major protocol deviations prior to database lock and final unblinding. Major deviations are those deviations that significantly impact efficacy and result in the removal of patients from the PP Analysis Set. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Study drug compliance $\leq 80\%$ or $\geq 105\%$
- Use of prohibited therapies
- Incorrect treatment

Major protocol deviations will be summarized by cohort and deviation category. Protocol deviations will be listed for the ITT Analysis Set by assigned treatment.

6.4. Demographic, Baseline Characteristics, and Disease History

Patient demographics and baseline characteristics will be summarized by cohort and overall for the ITT Analysis Set by assigned treatment. These will include age, sex (Male, Female), gender (Male, Female, Transgender Female to Male, Transgender Male to Female, Not willing to respond, Other), childbearing potential for females (Yes, No), ethnicity (Hispanic or Latino, Not Hispanic or Latino), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander), height (cm), weight (kg), and body mass index (BMI; kg/m^2). BMI will be calculated as (weight in kg) / (height in m)^2.

Demographics and baseline characteristics will be listed for the ITT Analysis Set.

Disease history will be summarized by cohort and overall for the ITT Analysis Set by assigned treatment. These will include frequency of pain in the last 3 months, how often this pain happens close in time to a bowel movement, how often when the pain happens are stools softer or harder than usual, how often when the pain happens are stools more frequent or less frequent than usual, how often the pain gets worse with menstrual bleeding for females, how often the pain gets worse after a meal, how often the pain limit usual activities, if the pain is continuous or not, if the pain has lasted for over 6 months or not, and abnormal stool condition. These are all collected from eCRF.

Disease history data will be listed for the ITT Analysis Set in by-patient listings.

6.5. Medical History

Medical history will be summarized by Medical Dictionary for Medical Affairs (MedDRA) system organ class (SOC) and preferred term (PT). Sort order will be by descending overall incidence of SOC followed by descending incidence of overall PT. Medical history will be summarized and presented by cohort and overall for the SAF by actual treatment.

Medical history will be presented in data listings for the SAF Analysis Set. Medical history coded terms will be provided, including the SOC and PT per MedDRA. The data management plan specifies the version of MedDRA used for medical history coding.

6.6. Prior and Concomitant Medication

The incidence of medication use will be summarized by the World Health Organization (WHO) Drug Dictionary Anatomic Therapeutic Chemical (ATC) Level 2 classification (i.e., therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study drug (i.e., taken exclusively during the pre-treatment period). Concomitant medications are those which have been identified to have been taken at any point during the on-treatment period, including medications which started prior to first dose of study drug but are ongoing at first dose. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Appendix 1 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented by cohort and overall for the SAF Analysis Set by actual treatment. Prior medications will be summarized in a similar manner.

During the treatment period, patients may also take protocol-permitted loperamide rescue medication for the acute treatment of uncontrolled diarrhea. Rescue medications will be summarized in a similar manner as the concomitant medication.

All prior and concomitant data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for the SAF. Prior and subsequent medications will be flagged by unique identifiers in the listing. Loperamide rescue medication use will also be flagged in the listing.

6.7. Rescue Medication

Rescue medication usage (loperamide) will be summarized by cohort and overall for all patients in the SAF Analysis Set. The time periods to be summarized will include the run-in period prior to Baseline/Day 1, the treatment period (Baseline/ Day1 through Day 29), and the post-treatment period (after Day 29 through Day 43). Categorical summaries will include the incidence of patients with rescue medication use at any time during the associated time period, as well as, the number of total doses of rescue medication used (0 doses, 1 dose, 2 doses, 3 doses, 4 doses, 5 doses, 6 doses, 7 doses, 8 doses, 9 doses, 10 or more doses) during the associated time period. The total number of doses of rescue medication also will be summarized as a continuous measure during the associated time period.

6.8. Study Drug Exposure and Compliance

Exposure to study drug will be summarized by cohort and overall for all patients in the SAF Analysis Set.

Patients will be randomized to receive for a total of 4 weeks BOS-589 at 1 of 2 dose levels or Placebo bid, as described in the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for the 200 mg BOS-589 bid orally
- Cohort 2 (Low Dose): 2 x 25 mg tablets for the 50 mg BOS-589 bid orally
- Cohort 3 (Placebo): 2 x visually matched Placebo tablets bid orally

The duration of exposure to study drug is calculated as date of last dose of study drug taken – date of first dose of study drug taken +1.

The planned total number of tablets during exposure is defined as duration of exposure x 4.

The actual total number of tablets taken during exposure is calculated as the total number of tablets given – the total number of tables returned, both collected in the eCRF.

The average actual daily number of tablets taken is calculated as the actual total number of tablets taken during exposure / duration of exposure to study drug.

The drug compliance is calculated as the actual total number of tablets taken during exposure / planned total number of tablets taken during exposure x 100%.

The duration of exposure to study drug, planned and actual total number of tablets during exposure, average actual daily number of tablets taken, and drug compliance will be summarized. Drug compliance categories of >=105%, >100% to <105%, >90% to <=100%, >80% to <=90%, >70% to <=80%, and <=70% will also be presented.

A by-patient listing of exposure and compliance measures will be produced for SAF Analysis Set.

7. EFFICACY

Efficacy analyses will be conducted using the ITT and PP Analysis Sets.

All available efficacy data will be provided in by-patient data listings on all ITT patients.

The efficacy objectives of the study are to evaluate the effect of BOS-589, relative to Placebo after 4 weeks of therapy, on pain, defecation, and key IBS-D related signs and symptoms.

During the 4 weeks of the double-blind treatment phase, patients will be required to access their electronic diary every day, preferably at the same time each day, to record IBS-D symptom data and information related to their bowel function and rescue medication use.

For all efficacy endpoints, the Day 29 visit change from baseline is the main timepoint of interest. However, unless stated otherwise, data will be presented for each planned visit (Baseline/Day 1, Day 8, Day 15, Day 22, Day 29, and Day 43).

For efficacy endpoints collected daily, the baseline score will be the average of non-missing scores reported for the last 7 days prior to the date where eligibility and run-in criteria pertaining to IBS-D symptom and bowel function have been met. The Day 29 visit score will be the average of non-missing scores reported during the 7 days prior to the Day 29 visit. The Day 8 visit, Day 15 visit, Day 22 visit, and Day 43 visit scores are calculated in a similar manner. If there are missing daily scores on any day, the visit score will be averaged from all the non-missing values during the 7 days prior to the visit. No imputation of the missing values will be made. For the Day 29 calculation only, a window of +/- 14 days relative to actual calculated day 29 from dosing will be applied such that any Day 29 visit occurring outside that window will be excluded.

7.1. Primary Efficacy Endpoint and Analyses

7.1.1. Worst Abdominal Pain

The primary efficacy endpoint will be defined as a change in the 24-hour WAP score at the Day 29 visit compared to baseline (averaged over the week prior to the visit). The 24-hour WAP score will be a patient-reported score of their WAP in the past 24 hours on a 0 to 10 scale, where 0 corresponds to no pain and 10 corresponds to worst imaginable pain. Patients will be asked to rate their WAP in the past 24 hours every day for the entire Run-in period and the 4-week double-blind treatment period.

This is a proof-of-principle study in which a hierarchical hypothesis test will be employed for the primary efficacy endpoint. The mean change in 24-hour WAP at the Day 29 visit from baseline (averaged over the week prior to the visit) will be measured as a continuous variable and compared between cohorts in the following testing order using a t-test of equal variances and a two-tailed alpha of 0.05:

Hypothesis 1:

Active Treatment (Cohort 1 + Cohort 2) versus Placebo (Cohort 3);

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

The null hypothesis will be that there is no difference in the mean change in 24-hour WAP score at the Day 29 visit from Baseline (averaged over the week prior to the visit) between active treatment and Placebo. The alternative hypothesis will be that there is a difference in the mean change in 24-hour WAP score at the Day 29 visit from Baseline (averaged over the week prior to the visit) between active treatment and Placebo. The test statistic is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

$\bar{x}_1$ and $\bar{x}_2$ are the sample means of change at the Day 29 visit from Baseline in 24-hour WAP score (averaged over the week prior to the visit), s^2 is the pooled sample variance assuming equal variance, n_1 and n_2 are the sample sizes, and the degree of freedom is $n_1 + n_2 - 2$.

If statistically significant at $P < 0.05$, then proceed to Hypothesis 2.

Hypothesis 2:

Cohort 1 (High Dose) versus Cohort 3 (Placebo);

If statistically significant at $P < 0.05$, then proceed to Hypothesis 3.

Hypothesis 3:

Cohort 2 (Low Dose) versus Cohort 3 (Placebo);

If statistically significant at $P < 0.05$ then, proceed to Hypothesis 4.

Hypothesis 4:

Cohort 1 (High Dose) versus Cohort 2 (Low Dose).

Methodology for hypotheses 2 to 4will be similar to that described for hypothesis 1.

The patient-reported 24-hour WAP scores and change from baseline scores (averaged over the week prior to each visit) will be summarized descriptively by cohort for each scheduled visit. Box plots of the 24-hour WAP scores (averaged over the week prior to each visit) will also be presented by cohort at each scheduled visit. Line plots of 24-hour WAP change from baseline scores (averaged over the week prior to each visit) will also be presented by cohort at each scheduled visit.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of worst abdominal pain will be repeated as a sensitivity analysis on this subgroup.

7.2. Secondary Efficacy Endpoints and Analyses

7.2.1. Stool Consistency

Patients will be asked to record daily stool consistency according to the BSFS most representative stool consistency in the past 24 hours and the worst stool consistency (defined as the loosest stool with the highest BSFS score) in the past 24 hours. The patient-reported BSFS consistency score, based on a 1 to 7 scale where 1 corresponds to a hard stool and 7 corresponds to watery diarrhea (10), can be found in Appendix 2.

The patient-reported stool consistency scores and change from baseline scores (averaged over the week prior to each visit) will be summarized descriptively by cohort for each scheduled timepoint. Pooled active treatment data will be compared to Placebo for the analyses, and by cohort, using a t-test of equal variances and a two-tailed alpha of 0.05 specified in section 7.1.1 for reference. Box plots of the BSFS scores (averaged over the week prior to each visit) will also be presented by cohort for each scheduled visit. Line plots of the BSFS change from baseline values (averaged over the week prior to each visit) will also be presented by cohort for each scheduled visit. Analyses for the most representative stool consistency in the past 24 hours and worst (loosest) stool consistency in the past 24 hours will both be presented.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of BSFS most representative consistency and BSFS worst consistency will be repeated as sensitivity analyses on this subgroup as described in section 7.1.1.

7.2.2. Stool Frequency

Patients will be asked to record stool frequency based on the total number of spontaneous bowel movements in the past 24 hours. Bowel movement frequency and change from baseline (averaged over the week prior to each visit) will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of stool frequency will be repeated as a sensitivity analysis on this subgroup as described in section 7.1.1.

7.2.3. Irritable Bowel Syndrome Severity Score

IBS-SS will be collected at only at the baseline, Day 29, and Day 43 visits. Patients will be asked to complete 5 questions regarding the severity of their IBS. Each of the 5 questions generate a maximum score of 100, leading to a total possible score of 500. IBS-SS will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1 (except only for scheduled visits of Baseline/Day 1, Day 29, and Day 43). Summaries performed by visit will include responses to the individual questions, as well as the total scores.

7.2.4. Irritable Bowel Syndrome Global Score

Patients will be asked to record daily their overall IBS-GS in the prior 24 hours. The patient-reported daily IBS-GS is based on a 0 to 4 scale where:

- 0 corresponds to no symptoms;
- 1 corresponds to mild symptoms;
- 2 corresponds to moderate symptoms;
- 3 corresponds to severe symptoms;
- 4 corresponds to very severe symptoms.

IBS-GS will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1.

7.3. Exploratory Efficacy Endpoints and Analyses

7.3.1. Weekly Responder

Weekly responders will be summarized as an exploratory efficacy endpoint.

A patient is categorized as a weekly responder if the patient is a weekly responder in both pain intensity and stool consistency (page 8 of FDA guidance document on IBS).

- An Abdominal Pain Intensity Weekly Responder is defined as a patient who experiences a decrease in the weekly average of the WAP in the past 24 hours score of at least 30 percent compared with weekly average at baseline, i.e., (weekly average of baseline – weekly average of a post-baseline visit) / weekly average of baseline >= 0.3
- A Stool Consistency Weekly Responder is defined as a patient who experiences a 50 percent or greater reduction in the number of days per week with at least one stool that has a consistency of Type 6 or 7 compared with weekly average at baseline, i.e., (weekly average of the number of Type 6 or 7 BSFS stool of baseline – weekly average of the number of Type 6 or 7 BSFS stool of a post-baseline visit) / weekly average of the number of Type 6 or 7 BSFS stool of baseline >= 0.5

The number and percentage of patients who are abdominal pain intensity weekly responders, stool consistency weekly responder, and composite weekly responders in both will be summarized by cohort at each scheduled visit. A chi-square test will be used, as a reference, to test whether there is a difference by between each active dose level and Placebo, both active dose levels and Placebo, and between the two active dose levels. Fisher's exact test will be used if the expected values in any of the cells of the contingency table are below 5.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of weekly responders will be repeated as a sensitivity analysis on this subgroup.

8. PHARMACOKINETICS / PHARMACODYNAMICS

PK/PD analyses are covered under a separate SAP.

9. SAFETY

Unless otherwise specified, all safety analysis reporting will be based on the SAF Analysis Dataset. For safety analyses, Day 29 represents measures taken at end of treatment/early termination while Day 43 represents measures taken at the end of study.

9.1. Adverse Events

AEs starting after the informed consent date until the Day 43/End of Study Follow-up visit will be recorded. AE assessments will include severity (mild, moderate, severe), relationship to study drug (related or not related), and seriousness. AEs with a missing severity will be classified as severe. AEs with a missing causality will be classified as related. AEs with missing seriousness will be classified as serious.

Any event reported on the CRF that occurs on or after the initiation of study drug is defined as a treatment-emergent adverse event (TEAE). Guidelines for categorizing an event as treatment-emergent based on start and stop dates and times are in Appendix 1. Additionally, an AE that is reported to have started on Day 1 without an associated onset time will be assumed to have occurred after the initiation of study drug. Hence, AEs occurring on Day 1 with no associated onset time will be assumed to be treatment emergent. SAEs associated with a protocol specified procedure and occurring after the time of consent but before administration of first dose of study drug will be defined as non-treatment emergent AEs (NTEAEs) and will be excluded from summary tables but will appear in listings flagged as NTEAE.

On a summary table, the number and percentage of patients in each of the following categories will be summarized by cohort following the cohort display in section 5.9:

- Any TEAE
- Any TEAEs related to study drug
- Any moderate or severe TEAEs
- Any moderate or severe TEAEs related to study drug
- Any treatment-emergent SAEs
- Any treatment-emergent SAEs related to study drug
- Any TEAEs causing discontinuation of study drug
- Any TEAEs with outcome of death
- Any TEAEs leading to withdrawal from the study

AEs will be coded to a SOC and PT using MedDRA (the version used will be defined in the study data management plan).

For summaries of TEAEs by SOC and PT, a patient will be counted once within an SOC, even if the patient experienced more than one TEAE within a specific SOC (likewise for PT). TEAEs will be ordered by descending frequency for overall SOC incidence followed by descending frequency of overall PT incidence. Summaries of TEAEs by SOC and PT will include the following:

- All TEAEs
- All TEAEs by severity
- All moderate or severe TEAEs
- All TEAEs related to study drug
- All treatment-emergent SAEs

- All treatment-emergent SAEs related to study drug
- All TEAEs causing discontinuation of study drug
- All TEAEs with outcome of death
- All TEAEs leading to withdrawal from the study

All AEs, including the verbatim and coded terms, will be listed for all enrolled patients. Additional separate listings will be provided for all AEs with outcome of death, all AEs causing discontinuation of study drug, all SAEs, and all AEs leading to withdrawal from the study.

9.2. Clinical Laboratory Evaluations

Laboratory tests identified in Appendix 1 in the protocol will be performed at times defined in the Schedule of Activities, Section 1.3 of the protocol. Clinical chemistry, hematology, and coagulation parameters will be reported based on the International System of Units (SI). All abnormal laboratory results will be evaluated by the Investigator as either clinically significant or not clinically significant.

Observed values and changes from baseline for hematology, clinical chemistry, urinalysis, and inflammatory biomarker laboratory evaluations will be summarized by cohort at each scheduled visit (Baseline/Day 1, Day 15, Day 29/Early Termination, and Day 43/End of Study) and most extreme change.

Box plots will be produced by visit and treatment for the recorded values of Erythrocyte Sedimentation Rate, C-reactive Protein, and Fecal Calprotectin. As well, line plots of the mean change from baseline of these three tests at each visit will be produced by treatment.

The number and percentage of patients experiencing treatment-emergent graded toxicities will be summarized by cohort and severity grade per the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.3. For each parameter, a patient will only be counted once using the worst-case value among all post-baseline timepoints. Laboratory toxicity will be summarized by incidence of CTCAE grade 3 and CTCAE grade 4 values for measures with a corresponding CTCAE grade. For measures without a corresponding CTCAE grade, incidence of 2.5 to < 5.0 X upper or lower limit of normal and ≥ 5.0 X upper or lower limit of normal cohort by cohort. Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with at least one post-baseline assessment for the laboratory parameter being displayed. Separate summary counts will be prepared for parameters with bi-directional toxicity grading with "high/hyper" conditions depending on the maximum post-baseline value and the "low/hypo" conditions depending on the minimum post-baseline value.

All laboratory parameters will be provided in patient data listings for all enrolled patients.

9.3. Other Safety Evaluations

For electrocardiogram and vital signs results, Marked Abnormality will be defined in the following sections. Should a patient have a Markedly Abnormal result at baseline, only those

post-baseline results that are at least 1 grade more extreme than the baseline grade will be considered as a Marked Abnormality.

9.3.1. 12-Lead Electrocardiogram

12-lead electrocardiograms (ECGs) will be assessed as specified in the Schedule of Activities in Section 1.3 of the protocol. The following ECG parameters will be collected: PR interval (milliseconds [msec]), QRS interval (msec), QT interval (msec), QT interval corrected for heart rate (QTc) (msec), and Ventricular Rate (beats/min). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values and changes from baseline for ECG parameters will be summarized at each scheduled visit (Screening and Day 29/ Early Termination) by cohort and overall.

Observed and change from baseline values on Day 29 in QTc status categorized as markedly abnormal (defined in the table below) or not will be summarized.

QTc Measure	**Markedly Abnormal Criteria**
Observed	450–480 msec, inclusive [CTCAE grade 1] 481–500 msec, inclusive [CTCAE grade 2] > 500 msec [CTCAE grade 3 or higher]
Change from Baseline	31–60 msec, inclusive, increase from baseline >60 msec increase from baseline

If triplicate measures were taken at a single time point, the average will be summarized.
All ECG data will be presented in a by patient data listing for all enrolled patients. Measures meeting the markedly abnormal criteria will be flagged in the listing.

9.3.2. Physical Examinations

Physical examinations will be presented in patient data listings on all enrolled patients.

9.3.3. Vital Signs

Vital signs include: pulse rate (beats/min), respiratory rate (breaths/min); temperature (degrees Celsius [°C]); systolic and diastolic blood pressure (millimeters of mercury [mmHg]); height, weight (kg), and BMI (derived).

Observed values and changes from baseline for vital signs will be summarized at by cohort and overall for each scheduled visit (Baseline/Day 1, Day 8, Day 15, Day 22, Day 29, and Day 43) and most extreme change.

Summaries of pulse rate, temperature, systolic blood pressure, and diastolic blood pressure will be based on markedly abnormal criteria defined below:

Vital Sign	Markedly Abnormal Criteria
Pulse rate (beats/min)	< 60 beats/min > 100 beats/min
Temperature (°C)	≤ 35 °C ≥ 38 °C
Systolic blood pressure (mmHg)	120-139 mmHg, inclusive (CTCAE grade 1) 140–159 mmHg, inclusive (CTCAE grade 2) ≥ 160 mmHg (CTCAE Grade 3)
Diastolic blood pressure (mmHg)	80–89 mmHg, inclusive (CTCAE grade 1) 90–99 mmHg, inclusive (CTCAE grade 2) ≥ 100 mmHg (CTCAE grade 3)

Incidence of post-baseline markedly abnormal worst-case values will be presented. For pulse rate and temperature, both high and low values will be presented separately such that patients can be counted in both categories.

All vital signs data will be presented in patient data listings for all enrolled patients. Markedly abnormal vital sign values will be flagged as such in the vital signs listing.

10. REFERENCES

1. CCI [redacted]
2. FDA Guidance for Industry Irritable Bowel Syndrome — Clinical Evaluation of Drugs for Treatment 2012: https://www.fda.gov/media/78622/download

11. APPENDICES

APPENDIX1 Partial Date Conventions

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

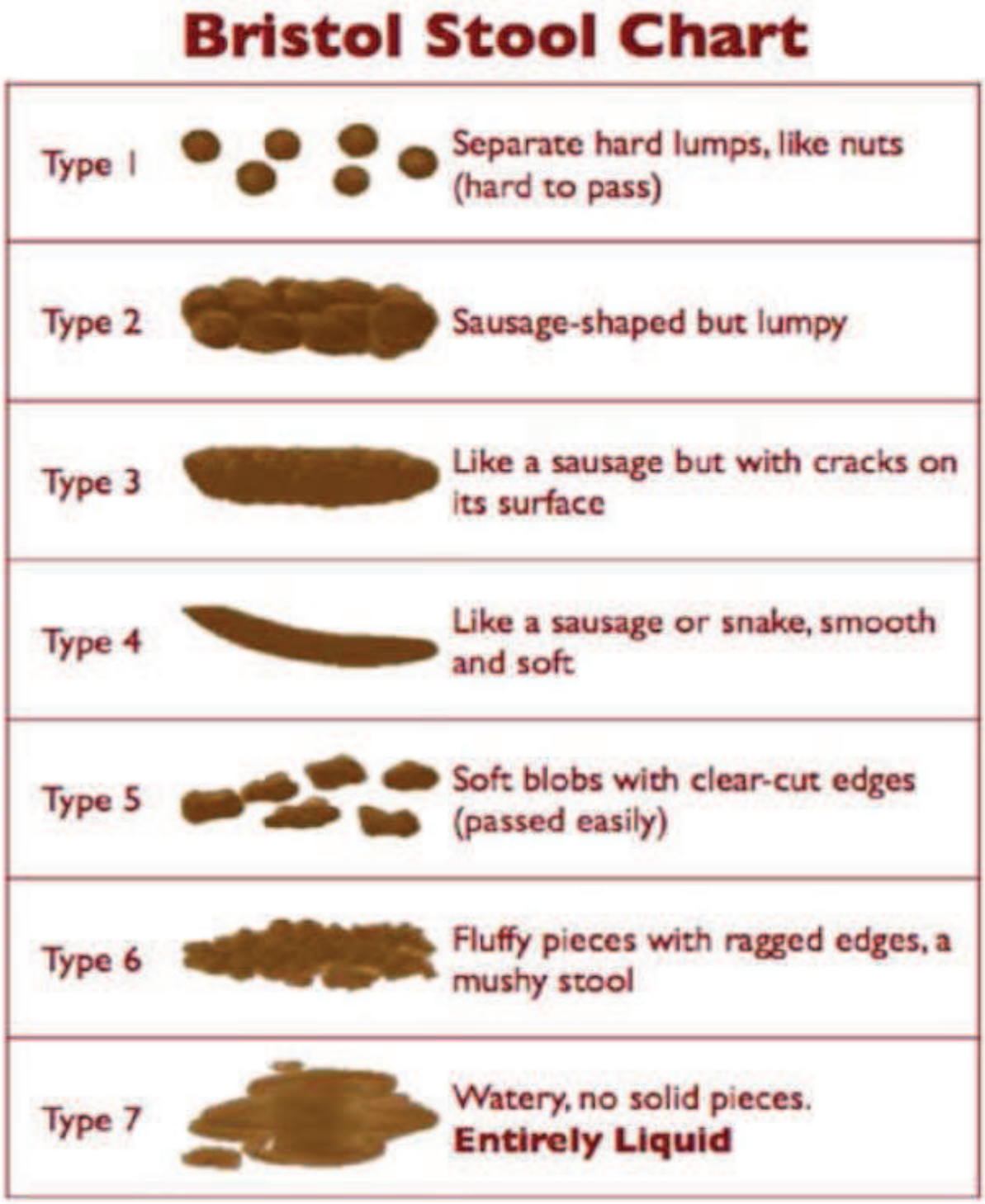
START DATE	STOP DATE	ACTION
Known	Known/Partial /Missing	If start date < study drug start date, then not TEAE If start date >= study drug start date and < end of treatment + 15 in days , then TEAE
Partial, but known components show that it cannot be on or after study drug start date	Known/Partial /Missing	Not TEAE
Partial, could be on or after study drug start date	Known	If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Missing	Assumed TEAE
Missing	Known	If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Missing	Assumed TEAE

ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS:

START DATE	STOP DATE	ACTION
Known	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Missing	If stop date is missing could never be assumed a prior medication If start date <= end of treatment + 15 in days, assign as concomitant
Partial	Known	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Partial	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown) and impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Missing	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown), then: If stop date is missing could never be assumed a prior medication If start date <= end of treatment + 15 in days, assign as concomitant
Missing	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant

START DATE	STOP DATE	ACTION
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant
	Missing	Assign as concomitant

APPENDIX2 The Bristol Stool Form Scale



A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE EFFICACY, SAFETY, AND TOLERABILITY OF BOS-589 IN THE TREATMENT OF PATIENTS WITH DIARRHEA-PREDOMINANT IRRITABLE BOWEL SYNDROME (IBS-D)

Statistical Analysis Plan

VERSION 1.1
DATE OF PLAN:

Mar 27, 2020

STUDY DRUG:
BOS-589

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Approval Signature Page: PPD Inc.

PPD

PPD

PPD

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD

4/7/2020

Date

PPD

4/7/2020

Date

Modification History

After approval of Version 1.0 of the Statistical Analysis Plan, subsequent versions should be documented below with a brief description of the change from the previous version as well as the rationale for the change.

Version, Date	Made by	Brief Description of Change and Rationale
Version 1.1, 27-MAR-2020	PPD	Baseline definition in section 7 was revised to be based on the 7 days prior to the date the noted eligibility criteria were met.

Contents

ABBREVIATIONS

°C	Degrees Celsius
AE	Adverse event
ATC	Anatomic Therapeutic Chemical
AUC	Area under the curve
AUC_{0-4}	Area under the concentration versus time curve from time zero to 4 hours postdose
AUC_{0-t}	Area under the concentration versus time curve from time zero to the last quantifiable concentration
bid	Twice daily
BMI	Body mass index
BSFS	Bristol Stool Form Score
cm	Centimeters
C_{max}	Maximum observed plasma concentration
C_{min}	Minimum observed plasma concentration
CTCAE	Common Terminology Criteria for Adverse Events
ECG	Electrocardiogram
eCRF	Electronic case report form
IBS	Irritable bowel syndrome
IBS-GS	Irritable Bowel Syndrome Global Scale
IBS-D	Irritable bowel syndrome-diarrhea
IBS SS	Irritable Bowel Syndrome Severity Score
ITT	Intention-To-Treat Analysis Set
IWRS	Interactive web response system
kg	Kilograms
m	Meters
MedDRA	Medical Dictionary for Medical Affairs
mg	Milligrams
mm	Millimeters
mmHg	Millimeters of mercury
msec	Milliseconds
NTEAE	Non-treatment emergent adverse event
PD	Progressive disease or Pharmacodynamics
PK	Pharmacokinetics
PPS	Per- Protocol Analysis Set
PT	Preferred Term
Q1	25^{th} Percentile/First quartile
Q3	75^{th} Percentile/Third quartile
SAE	Serious adverse events
SAF	Safety
SAP	Statistical analysis plan
SAS®	Statistical Analysis System
SD	Standard deviation
SI	International System of Units

SMC	Safety monitoring committee
SOC	System Organ Class
TEAE	Treatment-emergent adverse events
T_{max}	Time to maximum observed plasma concentration
WAP	Worst abdominal pain
WHO	World Health Organization

1. INTRODUCTION

The statistical analysis plan (SAP) details the planned analysis required to satisfy the Clinical Study Report of study number BOS-589-201: "A Phase 2a, Randomized, Double-blind, Placebo-controlled, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of BOS-589 in the Treatment of Patients with Diarrhea-predominant Irritable Bowel Syndrome (IBS-D)". The content of this SAP is based on the protocol amendment 2 (version 3.0) dated 12 Aug 2019.

Mock shells for tables, listings, and figures are in a separate document. Analyses for pharmacokinetics (PK) and pharmacodynamics (PD) will be governed by a separate SAP.

2. STUDY OBJECTIVES AND ENDPOINTS

OBJECTIVES	ENDPOINTS
Primary	
• To evaluate in patients with IBS-D the abdominal pain response to BOS-589 after 4 weeks of treatment, relative to Placebo.	• 24-hour worst abdominal pain scores (WAP) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the overall safety and tolerability of BOS-589 in the treatment of IBS-D during 4 weeks of treatment, relative to Placebo.	• Incidence of adverse events (AEs) serious adverse events (SAEs), discontinuations because of AEs, and any treatment-related severe AEs.
Secondary	
• To evaluate the treatment effect of BOS-589 on defecation after 4 weeks, relative to Placebo.	• Change in stool consistency, measured by the daily Bristol Stool Form Score (BSFS) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). • Change in stool frequency, measured by the total number of spontaneous bowel movements in 24 hours at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the treatment effect of BOS-589 on IBS-related signs and symptoms.	• Change in the IBS Severity Score (IBS-SS) at Day 29 compared to baseline. • Change in the IBS Global Scale (IBS-GS) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the steady-state PK of BOS-589.	• Maximum observed plasma concentration (C_{max}); time to C_{max} (T_{max}); minimum plasma concentration (C_{min}); area under the concentration versus time curve (AUC) from time zero to 4 hours postdose (AUC_{0-4}); AUC from time zero to the last quantifiable concentration (AUC_{0-t}).

Additional Objectives	Additional Endpoints
Exploratory (not included in protocol amendment 2 [version 3.0] dated 12 Aug 2019)	
• To explore the treatment effect of BOS-589 in terms of weekly responders after 4 weeks of treatment, relative to Placebo.	• Number of patients who are weekly responders, defined by both an abdominal pain intensity weekly responder (decrease in the weekly

Additional Objectives	Additional Endpoints
	average of the WAP in the past 24 hours score of at least 30%compared with baseline averaged over the week prior) and a stool consistency weekly responder (at least 50%reduction in the number of days per week with at least one stool that has a consistency of Type 6 or 7 compared with baseline averaged over the week prior), at each scheduled visit.

3. STUDY DESIGN

3.1. Study Design and Population

This study is a phase 2a, randomized, double-blind, Placebo-controlled, multicenter trial designed to provide the first proof-of-principle efficacy of BOS-589 in IBS-D patients, and to inform dose selection for subsequent development. The study will consist of a pre-treatment phase, a 4-week double-blind treatment phase, and a 2-week post-treatment follow-up period.

Pre-Treatment Phase: During the pre-treatment phase, patients will be evaluated for up to 5 weeks to assess eligibility. The pre-treatment phase will consist of initial screening assessments and a Run-in period.

After the initial screening assessments have been performed, eligible patients will enter a Run-in period of up to 3 weeks. During the Run-in period, the patients will complete an electronic diary to collect daily information related to their IBS-D symptoms, bowel function, and rescue medicine use, to confirm disease activity and diary compliance. Patients will also be requested to discontinue any prohibited medications during this phase of the study.

Upon completion of the Run-in period, patients will return to the study site to confirm eligibility for randomization into a 4-week double-blinded Treatment Phase.

Treatment Phase: Eligible patients will be randomized (1:1:1) into the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for 200 mg BOS-589 twice daily (bid) orally;
- Cohort 2 (Low Dose): 2 x 25 mg tablets for 50 mg BOS-589 bid orally;
- Cohort 3 (Placebo): 2 x visually matching Placebo oral tablets bid.

During the 4 weeks of double-blind treatment, patients will continue to record their daily IBS-D symptoms, bowel function, and rescue medicine use in the electronic diary, as described in Section 8.3 of the protocol.

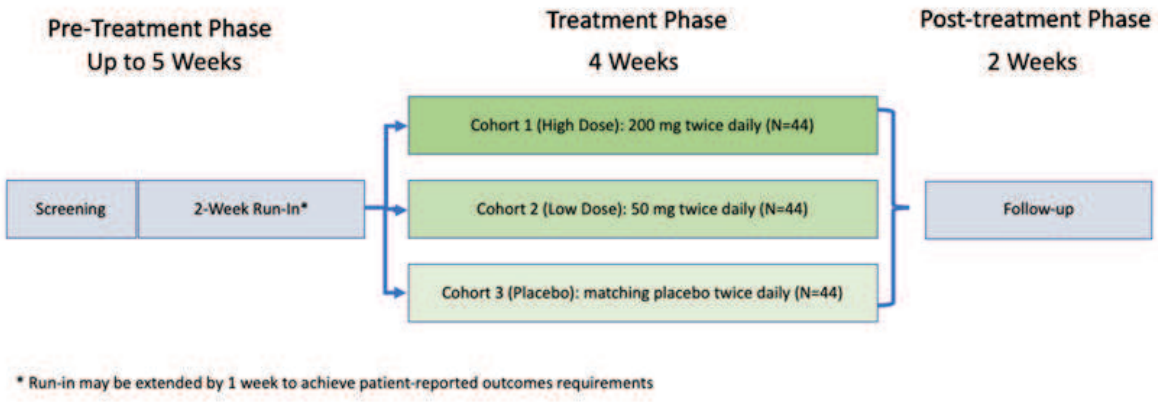
Post-treatment Phase: A 2-week post-treatment follow-up visit will occur for patients who complete the Treatment Phase. During the 2 weeks of follow-up, patients should continue to record their daily IBS-D symptoms, bowel function, and rescue medicine use in the electronic diary.

Patients who prematurely discontinue treatment should return to the study center to complete the early termination assessments as soon as possible after stopping the treatment.

Data analyses will occur after all patients in the trial have completed the last visit or procedure shown in the Schedule of Activities, Section 1.3 of the protocol.

Post-study Access to Therapy: No post-treatment access to therapy will be provided to patients randomized in the study.

Figure 1. Study Schematic



Further information on the study design, including eligibility criteria and a schedule of assessments, are detailed in the study protocol.

3.2. Changes to the Analyses from the Protocol

One exploratory efficacy endpoint was added in addition to the endpoints defined in the protocol.

- Summaries of weekly responders (defined in section 7.3~~7.4~~) will be summarized as an exploratory efficacy endpoint.

The wording of how randomization file is handled is clarified in the SAP by saying the file will be held by an independent unblinded Biostatistician, not the study statistician.

3.3. Randomization and Blinding

This is a randomized, double-blinded, Placebo-controlled study. Patients should be randomized as close as possible to the time of the planned first dose of study drug.

Randomization will be conducted via an interactive web response system (IWRS) in a 1:1:1 ratio. Eligible patients will be randomized (1:1:1) into the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for 200 mg BOS-589 twice daily (bid) orally;
- Cohort 2 (Low Dose): 2 x 25 mg tablets for 50 mg BOS-589 bid orally;
- Cohort 3 (Placebo): 2 x visually matching Placebo oral tablets bid.

Patients will be assigned to the cohorts using a permuted block randomization stratified by site (44 patients per cohort). The randomization schedule will be generated by an independent unblinded Biostatistician and will be transferred to the IWRS team for loading into the system. PPD blinded biostatistics and blinded programming team will remain blinded. The randomization file will be held by the independent unblinded Biostatistician until the end of the trial and when the database has been locked.

The sponsor, site staff, and patients will be blinded to the treatment assignment. BOS-589 and Placebo will be similar in appearance, and provided in white, opaque, high-density polyethylene bottles, each with a child-resistant closure. Each bottle will be labeled as per country requirement.

Only in a medical emergency will a patient's treatment assignment be unblinded, and this process will be performed and documented in the IWRS. Every effort should be made to contact the Medical Monitor to discuss unblinding prior to breaking the blind.

3.4. Sample Size Considerations

This is a phase 2a study in which a hierarchical hypothesis test will be employed for the primary efficacy endpoint. The primary endpoint is defined as a change in the 24-hour WAP score at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). Mean change from baseline will be measured as a continuous variable and compared between cohorts in a hierarchical testing order using a t-test of equal variances and a two-tailed alpha of 0.05.

The sample size was based on the first hypothesis test in the hierarchical plan, with approximately 80% power, two-sided alpha of 0.05, and a standard deviation (SD) of 3 points on the numeric rating scale to detect a 1.6 minimum change between cohorts (N = 120 patients). To account for attrition and a potential higher Placebo rate, a total of 132 patients will be randomized to one of 3 cohorts in a 1:1:1 ratio (44 patients per cohort).

3.5. Safety Monitoring Committee

There is no safety monitoring committee reviewing data for this study.

3.6. Interim Analysis

No formal efficacy interim analysis is planned for the purposes of statistically comparing the data prior to the completion of the study.

3.7. Timing of Analyses

Formal hypothesis testing will occur at the end of the study and post database lock. There are no planned Data, Safety, or Adjudication Committees for this study and no planned statistical interim analyses.

4. DATA ANALYSIS CONSIDERATIONS

Unless noted otherwise, the following statistical considerations will be applied:

- All analyses will be conducted based on Statistical Analysis System (SAS®) software version 9.4 or higher.
- Data for all sites/centers will be pooled for analysis.
- All data in the database except birth date will be presented in by-patient data listings.
- Listings will be sorted by cohort, patient identification number (concatenated site and patient number), and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.
- Continuous data will be summarized by cohort (or last assigned dose level, as applicable) based on n, mean, median, standard deviation (SD), 25th percentile/first quartile (Q1), 75th percentile/third quartile (Q3), minimum value, and maximum value.
- Categorical data will be summarized by cohort using n and percentage based on the number of non-missing values.
    - The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
    - Counts of zero will be presented without percentages.
- Statistics will be presented in the summary tables based on the following:
    - Mean, Median, Q1, and Q3: one additional decimal place to that reported for Minimum and Maximum
    - SD: two additional decimal places than the Minimum and Maximum

    - Percentages: reported to one decimal place
    - P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.
- The number and percentage will be presented in the form of XX (XX%). Percentage values of less than 1% will be presented as XX (<1%).
- Statistical inference will be based on a 5% significance level (i.e., 95% confidence intervals [CIs] will be produced). CIs will be produced only in cases with sufficient sample size (i.e., n => 5).
- Change from baseline = value (post-baseline visit) – value (baseline or last visit prior to treatment).
- All data up to the time of study completion/withdrawal from the study will be included in the analysis, regardless of duration of treatment.
- Numbering for data displays will be based on the International Council for Harmonization E3.
- The protocol uses the terms cohort, treatment group, and treatment arm interchangeably. For consistency in this document, the term cohort will be used.

4.1. Stratification and Covariates

There are no formal plans for analysis stratification or use of covariates.

4.2. Evaluation of Subgroups

No subgroup analyses are powered for the study. However, analyses for the primary and secondary endpoints will be analyzed by sex, and fecal calprotectin < 50 mcg/g at screening, and possibly other subgroups based on emerging data.

4.3. Multiple Comparisons and Multiplicity

A hierarchical hypothesis test will be employed for the primary efficacy endpoint, in the following pre-defined testing order using a t-test of equal variances and a two-tailed alpha of 0.05:

Hypothesis 1:

Active Treatment (Cohort 1 + Cohort 2) versus Placebo (Cohort 3);

If statistically significant at P < 0.05, then proceed to Hypothesis 2.

Hypothesis 2:

Cohort 1 (High Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05, then proceed to Hypothesis 3.

Hypothesis 3:

Cohort 2 (Low Dose) versus Cohort 3 (Placebo);

If statistically significant at $P < 0.05$ then, proceed to Hypothesis 4.

Hypothesis 4:

Cohort 1 (High Dose) versus Cohort 2 (Low Dose).

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Assigned and Actual Treatment

Patients will be randomized via IWRS in a 1:1:1 ratio to receive BOS-589 200 mg, BOS-589 50 mg, or Placebo. Patients will be assigned to the cohorts using a permuted block randomization stratified by site (44 patients per group).

The assigned treatment is determined by a randomization schedule. The randomization schedule will be generated and will be transferred to the IWRS team for loading into the system. The actual treatment given to individual patient is the actual treatment.

All safety analyses will be based on the actual treatment. Study patient data and all efficacy analyses will be based on the assigned treatment unless specified otherwise.

5.2. Reference Dates

Reference dates are defined as follows:

- Screening date is defined as the electronic case report form (eCRF) provided visit date on which a patient was screened for trial entry.
- Informed consent date is defined as the eCRF provided date of informed consent from the informed consent page.
- Run-in period start date is defined as earliest date of eCRF provided visit date from the Run-in period page.
- Run-in period end date is defined as eCRF provided date of randomization -1 from the randomization page.
- Treatment start date is defined as the date of first dose of any study drug.
- Treatment end date is defined as the date of last dose of any study drug.
- Age will come directly from the eCRF which is integrated from IWRS based on the informed consent date.
- Safety data, such as AEs and laboratory assessments, will use the treatment start date as a reference date.
- Efficacy data will use the treatment start date as a reference date.

- Study day will be based on treatment start date as a reference date. Study Day 1 will be the day patients start the first dose of study drug.

5.3. Study Day and Duration Variables

Reference date calculations will generally be defined as the following, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest >= reference date;
- otherwise, date of interest – reference date.

If either date is missing, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.6.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – informed consent date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug. Patients still receiving ongoing treatment or participating in study follow up at the time of analysis will use imputed treatment end and end of study dates as described in Section 5.7~~5.6~~.

Durations reported in days will be calculated as measure date – reference date + 1. When reporting in months the result will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.4. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date. Treatment end date is defined as the date of last dose of study drug.
- Post-treatment is defined as the period of time following the on-treatment period.

5.5. Baseline and Post-Baseline Changes

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value (including unscheduled assessments) collected prior to or on the treatment start date/time for measures collected prior to the treatment.
- Post-baseline values will be those collected after the date/time of first dose of study drug.

- Change from baseline is defined as: value – baseline value.
- Percentage change from baseline is defined as: (value – baseline value)/baseline value x 100%.
- Maximum change will be based on the change from baseline value to the maximum value for all post-baseline assessments collected during the on-treatment period, scheduled or unscheduled, using the formulas above.
- Minimum change will be based on the change from baseline value to the minimum value for all post-baseline assessments collected during the on-treatment period, scheduled or unscheduled, using the formulas above.

5.6. Imputation of Partial Dates

Appendix 1 details partial date conventions for AEs and medications.

Treatment End Date

Missing treatment end dates will not be imputed at the end of the study. However, due to ongoing reporting needs, treatment end date will be imputed as earliest of the data cutoff date, date of death, or last treatment date recorded on the exposure eCRF.

End of Study Date

Missing end of study dates will not necessarily be imputed at the end of the study. However, due to ongoing reporting needs, end of study dates will be imputed as the earliest of the data cutoff date, date of death, or last date recorded on the eCRF.

5.7. Multiple Assessments and Visit Windows

If there are multiple assessments at a scheduled visit, then the last assessment will be used in the analysis.

Nominal visits (e.g., those identified by the study eCRF) will be the basis of summarization and statistical analysis; visit windows for each scheduled visit can be found in the Schedule of Activities from the protocol. Unscheduled data may be included in summaries of baseline and maximum/minimum change values; summaries of specific abnormalities any time post-baseline; and patient data listings.

5.8. Lost to Follow Up or Lapse of Adequate Assessments

A patient will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site staff. Missing evaluations will not be imputed.

5.9. Cohort Display

Cohorts will be displayed with the following columns. See the mock shells for additional details and a visual representation of the cohort display.

- 200 mg BOS-589 bid
- 50 mg BOS-589 bid
- Active Treatment (Combined 200 mg and 50 mg Cohorts)
- Placebo bid
- Overall (not needed for efficacy outputs)

For primary efficacy analyses testing, cohort significant p-values will be displayed in the hierarchical order below. Following the first p-value, each subsequent p-value will only be displayed if the prior p-value is < 0.05.

- Active Treatment (Combined 200 mg and 50 mg Cohorts) vs. Placebo
- 200 mg BOS-589 bid vs. Placebo
- 50 mg BOS-589 bid vs. Placebo
- 200 mg BOS-589 bid vs. 50 mg BOS-589 bid

5.10. Missing Data

Missing data imputations for efficacy endpoints are described in Section 7; AE and concomitant medication, ongoing end of treatment, and ongoing end of study are described in Section 5.6. Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Sets

- The Screened Analysis Set will include all patients who have signed an informed consent document.
- The Intention-to-Treat (ITT) Analysis Set will include all randomized patients.
- The Per Protocol (PP) Analysis Set will include all patients in the ITT Analysis Set who received at least 1 dose of study drug and did not have any major protocol deviations.
- The Safety (SAF) Analysis Set will include all patients who received at least 1 dose of study drug.

All safety analyses will be performed in the SAF Analysis Set by the actual cohorts. All efficacy analyses will be performed in the ITT and PP Analysis Sets by the randomized cohorts. Summaries of the Screened Analysis Set will be performed by the randomized cohorts.

The criteria for inclusion in the PP Analysis Set will be finalized and documented before database is locked and study unblinding. Additional information on inclusion into this population can be found in Section 6.3.

6.2. Patient Disposition

A summary of patient enrollment by site will be presented by cohort and overall for the Screened Set. The overall patient enrollment status (screened [i.e., consented], screen failures, randomized) will be provided. Patients who sign the ICF and are randomized but do not receive the study drug may be replaced. Once randomized, patients who are withdrawn from therapy before completing 28 days of dosing or are discontinued from the study for any reason will not be replaced.

Disposition data will be summarized for the Screened Set. The number of patients in each analysis set (ITT, SAF, PPS), completed treatment, with ongoing treatment, completed the study, and still ongoing in the study, will be summarized. The number of patients who discontinued study drug including reasons, the duration on study drug, and duration on study will be summarized. The number of patients who discontinued the study including reasons, the duration on study drug, and duration on study will also be summarized. Data will be presented by cohort and overall by assigned treatment.

A by-patient listing of patient enrollment data as well as a by-patient listing of patient disposition data including reason for treatment and study discontinuation will be presented for the screened patients.

6.3. Protocol Deviations

Protocol deviations will be recorded in the clinical trial management system.

Protocol deviations will be identified and classified as major protocol deviations prior to database lock and final unblinding. Major deviations are those deviations that significantly impact efficacy and result in the removal of patients from the PP Analysis Set. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Study drug compliance $\leq 80\%$ or $\geq 105\%$
- Use of prohibited therapies
- Incorrect treatment

Major protocol deviations will be summarized by cohort and deviation category. Protocol deviations will be listed for the ITT Analysis Set by assigned treatment.

6.4. Demographic, Baseline Characteristics, and Disease History

Patient demographics and baseline characteristics will be summarized by cohort and overall for the ITT Analysis Set by assigned treatment. These will include age, sex (Male, Female), gender (Male, Female, Transgender Female to Male, Transgender Male to Female, Not willing to respond, Other), childbearing potential for females (Yes, No), ethnicity (Hispanic or Latino, Not Hispanic or Latino), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander), height (cm), weight (kg), and body mass index (BMI; kg/m^2). BMI will be calculated as (weight in kg) / (height in m)^2.

Demographics and baseline characteristics will be listed for the ITT Analysis Set.

Disease history will be summarized by cohort and overall for the ITT Analysis Set by assigned treatment. These will include frequency of pain in the last 3 months, how often this pain happens close in time to a bowel movement, how often when the pain happens are stools softer or harder than usual, how often when the pain happens are stools more frequent or less frequent than usual, how often the pain gets worse with menstrual bleeding for females, how often the pain gets worse after a meal, how often the pain limit usual activities, if the pain is continuous or not, if the pain has lasted for over 6 months or not, and abnormal stool condition. These are all collected from eCRF.

Disease history data will be listed for the ITT Analysis Set in by-patient listings.

6.5. Medical History

Medical history will be summarized by Medical Dictionary for Medical Affairs (MedDRA) system organ class (SOC) and preferred term (PT). Sort order will be by descending overall incidence of SOC followed by descending incidence of overall PT. Medical history will be summarized and presented by cohort and overall for the SAF by actual treatment.

Medical history will be presented in data listings for the SAF Analysis Set. Medical history coded terms will be provided, including the SOC and PT per MedDRA. The data management plan specifies the version of MedDRA used for medical history coding.

6.6. Prior and Concomitant Medication

The incidence of medication use will be summarized by the World Health Organization (WHO) Drug Dictionary Anatomic Therapeutic Chemical (ATC) Level 2 classification (i.e., therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study drug (i.e., taken exclusively during the pre-treatment period). Concomitant medications are those which have been identified to have been taken at any point during the on-treatment period, including medications which started prior to first dose of study drug but are ongoing at first dose. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Appendix 1 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented by cohort and overall for the SAF Analysis Set by actual treatment. Prior medications will be summarized in a similar manner.

During the treatment period, patients may also take protocol-permitted loperamide rescue medication for the acute treatment of uncontrolled diarrhea. Rescue medications will be summarized in a similar manner as the concomitant medication.

All prior and concomitant data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for the SAF. Prior and subsequent medications will be flagged by unique identifiers in the listing. Loperamide rescue medication use will also be flagged in the listing.

6.7. Rescue Medication

Rescue medication usage (loperamide) will be summarized by cohort and overall for all patients in the SAF Analysis Set. The time periods to be summarized will include the run-in period prior to Baseline/Day 1, the treatment period (Baseline/ Day1 through Day 29), and the post-treatment period (after Day 29 through Day 43). Categorical summaries will include the incidence of patients with rescue medication use at any time during the associated time period, as well as, the number of total doses of rescue medication used (0 doses, 1 dose, 2 doses, 3 doses, 4 doses, 5 doses, 6 doses, 7 doses, 8 doses, 9 doses, 10 or more doses) during the associated time period. The total number of doses of rescue medication also will be summarized as a continuous measure during the associated time period.

6.8. Study Drug Exposure and Compliance

Exposure to study drug will be summarized by cohort and overall for all patients in the SAF Analysis Set.

Patients will be randomized to receive for a total of 4 weeks BOS-589 at 1 of 2 dose levels or Placebo bid, as described in the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for the 200 mg BOS-589 bid orally
- Cohort 2 (Low Dose): 2 x 25 mg tablets for the 50 mg BOS-589 bid orally
- Cohort 3 (Placebo): 2 x visually matched Placebo tablets bid orally

The duration of exposure to study drug is calculated as date of last dose of study drug taken – date of first dose of study drug taken +1.

The planned total number of tablets during exposure is defined as duration of exposure x 4.

The actual total number of tablets taken during exposure is calculated as the total number of tablets given – the total number of tables returned, both collected in the eCRF.

The average actual daily number of tablets taken is calculated as the actual total number of tablets taken during exposure / duration of exposure to study drug.

The drug compliance is calculated as the actual total number of tablets taken during exposure / planned total number of tablets taken during exposure x 100%.

The duration of exposure to study drug, planned and actual total number of tablets during exposure, average actual daily number of tablets taken, and drug compliance will be summarized. Drug compliance categories of >=105%, >100% to <105%, >90% to <=100%, >80% to <=90%, >70% to <=80%, and <=70% will also be presented.

A by-patient listing of exposure and compliance measures will be produced for SAF Analysis Set.

7. EFFICACY

Efficacy analyses will be conducted using the ITT and PP Analysis Sets.

All available efficacy data will be provided in by-patient data listings on all ITT patients.

The efficacy objectives of the study are to evaluate the effect of BOS-589, relative to Placebo after 4 weeks of therapy, on pain, defecation, and key IBS-D related signs and symptoms.

During the 4 weeks of the double-blind treatment phase, patients will be required to access their electronic diary every day, preferably at the same time each day, to record IBS-D symptom data and information related to their bowel function and rescue medication use.

For all efficacy endpoints, the Day 29 change from baseline (with scores averaged over the week prior to each respective timepoint) is the main timepoint of interest. However, unless stated otherwise, data will be presented for each scheduled timepoint (Baseline/Day 1, Day 8, Day 15, Day 22, Day 29, and Day 43). The baseline score will be the average of non-missing scores reported for the last 7 days prior to ~~and inclusive of~~ the date where eligibility and run-in criteria pertaining to IBS-D symptom and bowel function have been met. The Day 29 score will be the average of non-missing scores reported from Day 22 to Day 28. The Day 8, Day 15, Day 22, and Day 43 timepoints are calculated in a similar manner. If there are missing scores on any day, the weekly average will be taken from all the non-missing values. No imputation of the missing values will be made.

7.1. Primary Efficacy Endpoint and Analyses

7.1.1. Worst Abdominal Pain

The primary efficacy endpoint will be defined as a change in the 24-hour WAP score at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). The 24-hour WAP score will be a patient-reported score of their WAP in the past 24 hours on a 0 to 10 scale, where 0 corresponds to no pain and 10 corresponds to worst imaginable pain. Patients will be asked to rate their WAP in the past 24 hours every day for the entire Run-in period and the 4-week double-blind treatment period.

This is a proof-of-principle study in which a hierarchical hypothesis test will be employed for the primary efficacy endpoint. The mean change in 24-hour WAP at Day 29 from baseline (averaged over the week prior to each respective timepoint) will be measured as a continuous variable and compared between cohorts in the following testing order using a t-test of equal variances and a two-tailed alpha of 0.05:

Hypothesis 1:

Active Treatment (Cohort 1 + Cohort 2) versus Placebo (Cohort 3);

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

The null hypothesis will be that there is no difference in the mean change in 24-hour WAP score at Day 29 from Baseline (averaged over the week prior to each respective timepoint) between active treatment and Placebo. The alternative hypothesis will be that there is a difference in the mean change in 24-hour WAP score at Day 29 from Baseline (averaged over the week prior to each respective timepoint) between active treatment and Placebo. The test statistic is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

$\bar{x}_1$ and $\bar{x}_2$ are the sample means of change at Day 29 from Baseline in 24-hour WAP score (averaged over the week prior to each respective timepoint), s^2 is the pooled sample variance assuming equal variance, n_1 and n_2 are the sample sizes, and the degree of freedom is $n_1 + n_2 - 2$.

If statistically significant at $P < 0.05$, then proceed to Hypothesis 2.

Hypothesis 2:

Cohort 1 (High Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05, then proceed to Hypothesis 3.

Hypothesis 3:

Cohort 2 (Low Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05 then, proceed to Hypothesis 4.

Hypothesis 4:

Cohort 1 (High Dose) versus Cohort 2 (Low Dose).

Methodology for hypotheses 2 to 4will be similar to that described for hypothesis 1.

The patient-reported 24-hour WAP scores and change from baseline scores (averaged over the week prior to each respective timepoint) will be summarized descriptively by cohort for each scheduled visit. Box plots of the 24-hour WAP scores (averaged over the week prior for each respective timepoint) will also be presented by cohort at each scheduled visit. Line plots of 24-hour WAP change from baseline scores (averaged over the week prior for each respective timepoint) will also be presented by cohort at each scheduled visit.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of worst abdominal pain will be repeated as a sensitivity analysis on this subgroup.

7.2. Secondary Efficacy Endpoints and Analyses

7.2.1. Stool Consistency

Patients will be asked to record daily stool consistency according to the BSFS most representative stool consistency in the past 24 hours and the worst stool consistency (defined as the loosest stool with the highest BSFS score) in the past 24 hours. The patient-reported BSFS consistency score, based on a 1 to 7 scale where 1 corresponds to a hard stool and 7 corresponds to watery diarrhea (10[redacted]), can be found in Appendix 2.

The patient-reported stool consistency scores and change from baseline scores (averaged over the week prior to each respective timepoint) will be summarized descriptively by cohort for each scheduled timepoint. Pooled active treatment data will be compared to Placebo for the analyses, and by cohort, using a t-test of equal variances and a two-tailed alpha of 0.05 specified in section 7.1.1 for reference. Box plots of the BSFS scores (averaged over the week prior to each respective timepoint) will also be presented by cohort for each scheduled visit. Line plots of the BSFS change from baseline values (averaged over the week prior to each respective timepoint) will also be presented by cohort for each scheduled visit. Analyses for the most representative stool consistency in the past 24 hours and worst (loosest) stool consistency in the past 24 hours will both be presented.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of BSFS most representative consistency and BSFS worst consistency will be repeated as sensitivity analyses on this subgroup as described in section 7.1.1.

7.2.2. Stool Frequency

Patients will be asked to record stool frequency based on the total number of spontaneous bowel movements in the past 24 hours. Bowel movement frequency and change from baseline (averaged over the week prior to each respective timepoint) will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of stool frequency will be repeated as a sensitivity analysis on this subgroup as described in section 7.1.1.

7.2.3. Irritable Bowel Syndrome Severity Score

IBS-SS will be collected at baseline and Day 29 only. Patients will be asked to complete 5 questions regarding the severity of their IBS. Each of the 5 questions generate a maximum score of 100, leading to a total possible score of 500. IBS-SS will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1 (except only for scheduled visits of Baseline/Day 1, Day 29, and Day 43). Summaries performed by visit will include responses to the individual questions, as well as the total scores.

7.2.4. Irritable Bowel Syndrome Global Score

Patients will be asked to record daily their overall IBS-GS in the prior 24 hours. The patient-reported daily IBS-GS is based on a 0 to 4 scale where:

- 0 corresponds to no symptoms;
- 1 corresponds to mild symptoms;
- 2 corresponds to moderate symptoms;
- 3 corresponds to severe symptoms;
- 4 corresponds to very severe symptoms.

IBS-GS will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1 (except only for scheduled visits of Baseline/Day 1 and Day 29).

7.3. Exploratory Efficacy Endpoints and Analyses

7.3.1. Weekly Responder

Weekly responders will be summarized as an exploratory efficacy endpoint.

A patient is categorized as a weekly responder if the patient is a weekly responder in both pain intensity and stool consistency (page 8 of FDA guidance document on IBS).

- An Abdominal Pain Intensity Weekly Responder is defined as a patient who experiences a decrease in the weekly average of the WAP in the past 24 hours score of at least 30 percent compared with weekly average at baseline, i.e., (weekly average of baseline – weekly average of a post-baseline visit) / weekly average of baseline >= 0.3
- A Stool Consistency Weekly Responder is defined as a patient who experiences a 50 percent or greater reduction in the number of days per week with at least one stool that has a consistency of Type 6 or 7 compared with weekly average at baseline, i.e., (weekly average of the number of Type 6 or 7 BSFS stool of baseline – weekly average of the number of Type 6 or 7 BSFS stool of a post-baseline visit) / weekly average of the number of Type 6 or 7 BSFS stool of baseline >= 0.5

The number and percentage of patients who are abdominal pain intensity weekly responders, stool consistency weekly responder, and weekly responders in both will be summarized by cohort at each scheduled visit. A chi-square test will be used, as a reference, to test whether there is a difference by between each active dose level and Placebo, both active dose levels and Placebo, and between the two active dose levels. Fisher's exact test will be used if the expected values in any of the cells of the contingency table are below 5.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of weekly responders will be repeated as a sensitivity analysis on this subgroup.

8. PHARMACOKINETICS / PHARMACODYNAMICS

PK/PD analyses are covered under a separate SAP.

9. SAFETY

Unless otherwise specified, all safety analysis reporting will be based on the SAF Analysis Dataset. For safety analyses, Day 29 represents measures taken at end of treatment/early termination while Day 43 represents measures taken at the end of study.

9.1. Adverse Events

AEs starting after the informed consent date until the Day 43/End of Study Follow-up visit will be recorded. AE assessments will include severity (mild, moderate, severe), relationship to study drug (related or not related), and seriousness. AEs with a missing severity will be classified as severe. AEs with a missing causality will be classified as related. AEs with missing seriousness will be classified as serious.

Any event reported on the CRF that occurs on or after the initiation of study drug is defined as a treatment-emergent adverse event (TEAE). Guidelines for categorizing an event as treatment-emergent based on start and stop dates and times are in Appendix 1. Additionally, an AE that is reported to have started on Day 1 without an associated onset time will be assumed to have occurred after the initiation of study drug. Hence, AEs occurring on Day 1 with no associated onset time will be assumed to be treatment emergent. SAEs associated with a protocol specified procedure and occurring after the time of consent but before administration of first dose of study drug will be defined as non-treatment emergent AEs (NTEAEs) and will be excluded from summary tables but will appear in listings flagged as NTEAE.

On a summary table, the number and percentage of patients in each of the following categories will be summarized by cohort following the cohort display in section 5.9:

- Any TEAE
- Any TEAEs related to study drug
- Any moderate or severe TEAEs
- Any moderate or severe TEAEs related to study drug
- Any treatment-emergent SAEs
- Any treatment-emergent SAEs related to study drug
- Any TEAEs causing discontinuation of study drug
- Any TEAEs with outcome of death
- Any TEAEs leading to withdrawal from the study

AEs will be coded to a SOC and PT using MedDRA (the version used will be defined in the study data management plan).

For summaries of TEAEs by SOC and PT, a patient will be counted once within an SOC, even if the patient experienced more than one TEAE within a specific SOC (likewise for PT). TEAEs will be ordered by descending frequency for overall SOC incidence followed by descending frequency of overall PT incidence. Summaries of TEAEs by SOC and PT will include the following:

- All TEAEs
- All TEAEs by severity
- All moderate or severe TEAEs
- All TEAEs related to study drug
- All treatment-emergent SAEs

- All treatment-emergent SAEs related to study drug
- All TEAEs causing discontinuation of study drug
- All TEAEs with outcome of death
- All TEAEs leading to withdrawal from the study

All AEs, including the verbatim and coded terms, will be listed for all enrolled patients. Additional separate listings will be provided for all AEs with outcome of death, all AEs causing discontinuation of study drug, all SAEs, and all AEs leading to withdrawal from the study.

9.2. Clinical Laboratory Evaluations

Laboratory tests identified in Appendix 1 in the protocol will be performed at times defined in the Schedule of Activities, Section 1.3 of the protocol. Clinical chemistry, hematology, and coagulation parameters will be reported based on the International System of Units (SI). All abnormal laboratory results will be evaluated by the Investigator as either clinically significant or not clinically significant.

Observed values and changes from baseline for hematology, clinical chemistry, urinalysis, and inflammatory biomarker laboratory evaluations will be summarized by cohort at each scheduled visit (Baseline/Day 1, Day 15, Day 29/Early Termination, and Day 43/End of Study) and most extreme change.

Box plots will be produced by visit and treatment for the recorded values of Erythrocyte Sedimentation Rate, C-reactive Protein, and Fecal Calprotectin. As well, line plots of the mean change from baseline of these three tests at each visit will be produced by treatment.

The number and percentage of patients experiencing treatment-emergent graded toxicities will be summarized by cohort and severity grade per the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.3. For each parameter, a patient will only be counted once using the worst-case value among all post-baseline timepoints. Laboratory toxicity will be summarized by incidence of CTCAE grade 3 and CTCAE grade 4 values for measures with a corresponding CTCAE grade. For measures without a corresponding CTCAE grade, incidence of 2.5 to < 5.0 X upper or lower limit of normal and ≥ 5.0 X upper or lower limit of normal cohort by cohort. Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with at least one post-baseline assessment for the laboratory parameter being displayed. Separate summary counts will be prepared for parameters with bi-directional toxicity grading with "high/hyper" conditions depending on the maximum post-baseline value and the "low/hypo" conditions depending on the minimum post-baseline value.

All laboratory parameters will be provided in patient data listings for all enrolled patients.

9.3. Other Safety Evaluations

For electrocardiogram and vital signs results, Marked Abnormality will be defined in the following sections. Should a patient have a Markedly Abnormal result at baseline, only those

post-baseline results that are at least 1 grade more extreme than the baseline grade will be considered as a Marked Abnormality.

9.3.1. 12-Lead Electrocardiogram

12-lead electrocardiograms (ECGs) will be assessed as specified in the Schedule of Activities in Section 1.3 of the protocol. The following ECG parameters will be collected: PR interval (milliseconds [msec]), QRS interval (msec), QT interval (msec), QT interval corrected for heart rate (QTc) (msec), and Ventricular Rate (beats/min). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values and changes from baseline for ECG parameters will be summarized at each scheduled visit (Screening and Day 29/ Early Termination) by cohort and overall.

Observed and change from baseline values on Day 29 in QTc status categorized as markedly abnormal (defined in the table below) or not will be summarized.

QTc Measure	**Markedly Abnormal Criteria**
Observed	450–480 msec, inclusive [CTCAE grade 1] 481–500 msec, inclusive [CTCAE grade 2] > 500 msec [CTCAE grade 3 or higher]
Change from Baseline	31–60 msec, inclusive, increase from baseline >60 msec increase from baseline

If triplicate measures were taken at a single time point, the average will be summarized.
All ECG data will be presented in a by patient data listing for all enrolled patients. Measures meeting the markedly abnormal criteria will be flagged in the listing.

9.3.2. Physical Examinations

Physical examinations will be presented in patient data listings on all enrolled patients.

9.3.3. Vital Signs

Vital signs include: pulse rate (beats/min), respiratory rate (breaths/min); temperature (degrees Celsius [°C]); systolic and diastolic blood pressure (millimeters of mercury [mmHg]); height, weight (kg), and BMI (derived).

Observed values and changes from baseline for vital signs will be summarized at by cohort and overall for each scheduled visit (Baseline/Day 1, Day 8, Day 15, Day 22, Day 29, and Day 43) and most extreme change.

Summaries of pulse rate, temperature, systolic blood pressure, and diastolic blood pressure will be based on markedly abnormal criteria defined below:

Vital Sign	Markedly Abnormal Criteria
Pulse rate (beats/min)	< 60 beats/min > 100 beats/min
Temperature (°C)	≤ 35 °C ≥ 38 °C
Systolic blood pressure (mmHg)	120-139 mmHg, inclusive (CTCAE grade 1) 140–159 mmHg, inclusive (CTCAE grade 2) ≥ 160 mmHg (CTCAE Grade 3)
Diastolic blood pressure (mmHg)	80–89 mmHg, inclusive (CTCAE grade 1) 90–99 mmHg, inclusive (CTCAE grade 2) ≥ 100 mmHg (CTCAE grade 3)

Incidence of post-baseline markedly abnormal worst-case values will be presented. For pulse rate and temperature, both high and low values will be presented separately such that patients can be counted in both categories.

All vital signs data will be presented in patient data listings for all enrolled patients. Markedly abnormal vital sign values will be flagged as such in the vital signs listing.

10. REFERENCES

1. CCI [redacted]
2. FDA Guidance for Industry Irritable Bowel Syndrome — Clinical Evaluation of Drugs for Treatment 2012: https://www.fda.gov/media/78622/download

11. APPENDICES

APPENDIX1 Partial Date Conventions

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

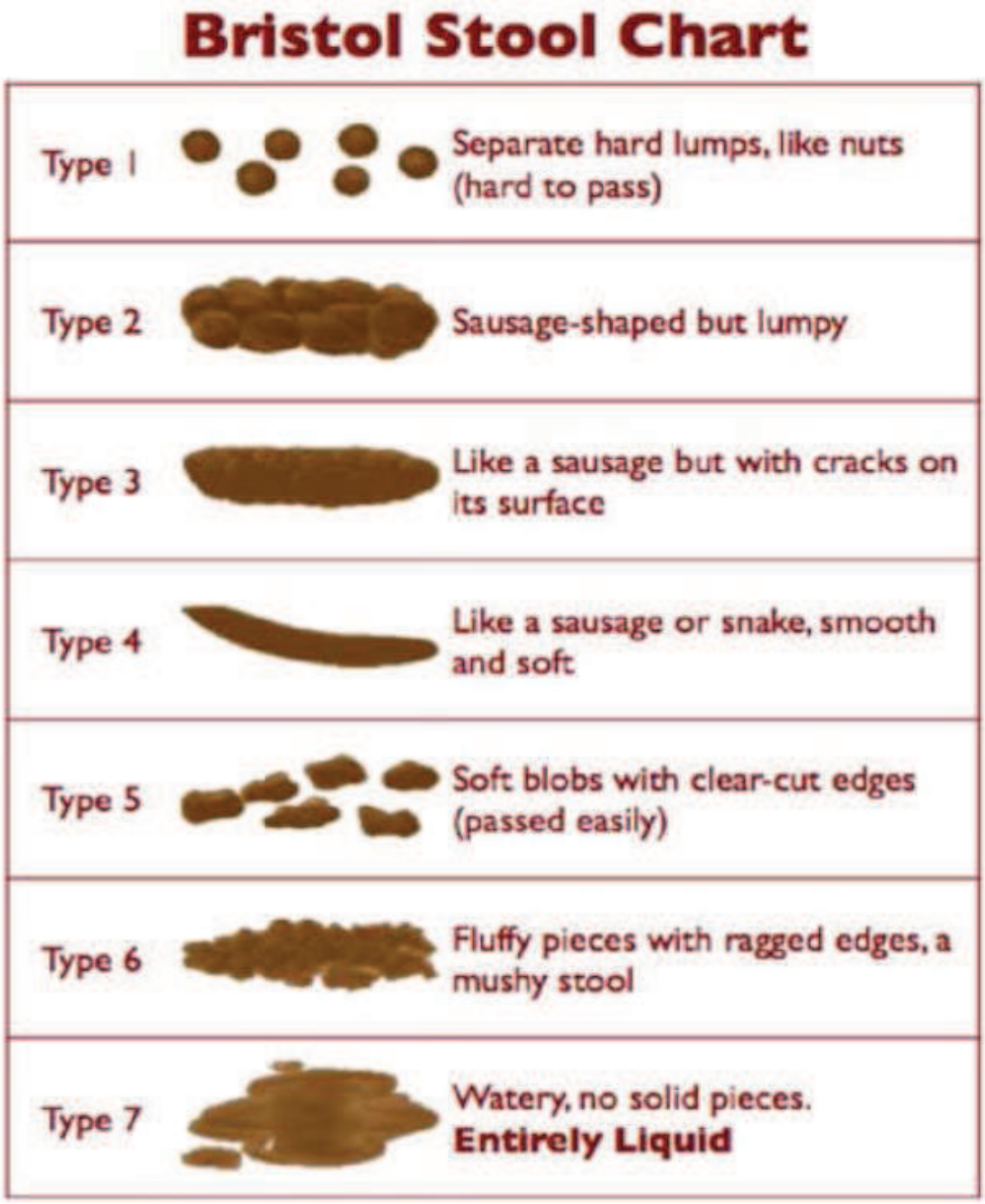
START DATE	STOP DATE	ACTION
Known	Known/Partial /Missing	If start date < study drug start date, then not TEAE If start date >= study drug start date and < end of treatment + 15 in days , then TEAE
Partial, but known components show that it cannot be on or after study drug start date	Known/Partial /Missing	Not TEAE
Partial, could be on or after study drug start date	Known	If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Missing	Assumed TEAE
Missing	Known	If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Missing	Assumed TEAE

ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS:

START DATE	STOP DATE	ACTION
Known	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Missing	If stop date is missing could never be assumed a prior medication If start date <= end of treatment + 15 in days, assign as concomitant
Partial	Known	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Partial	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown) and impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Missing	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown), then: If stop date is missing could never be assumed a prior medication If start date <= end of treatment + 15 in days, assign as concomitant
Missing	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant

START DATE	STOP DATE	ACTION
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant
	Missing	Assign as concomitant

APPENDIX2 The Bristol Stool Form Scale



A PHASE 2A, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED, MULTICENTER STUDY TO EVALUATE THE EFFICACY, SAFETY, AND TOLERABILITY OF BOS-589 IN THE TREATMENT OF PATIENTS WITH DIARRHEA-PREDOMINANT IRRITABLE BOWEL SYNDROME (IBS-D)

Statistical Analysis Plan

VERSION 1.0
DATE OF PLAN:

Jan 29, 2020

STUDY DRUG:
BOS-589

PREPARED FOR:

Boston Pharmaceuticals, Inc.

Approval Signature Page: PPD Inc.

PPD

PPD

PPD

Approval Signature Page: Boston Pharmaceuticals, Inc.

PPD

1/29/2020

Date

PPD

1/29/2020

Date

Modification History

After approval of Version 1.0 of the Statistical Analysis Plan, subsequent versions should be documented below with a brief description of the change from the previous version as well as the rationale for the change.

Version, Date	Made by	Brief Description of Change and Rationale

Contents

ABBREVIATIONS

°C	Degrees Celsius
AE	Adverse event
ATC	Anatomic Therapeutic Chemical
AUC	Area under the curve
AUC_{0-4}	Area under the concentration versus time curve from time zero to 4 hours postdose
AUC_{0-t}	Area under the concentration versus time curve from time zero to the last quantifiable concentration
bid	Twice daily
BMI	Body mass index
BSFS	Bristol Stool Form Score
cm	Centimeters
C_{max}	Maximum observed plasma concentration
C_{min}	Minimum observed plasma concentration
CTCAE	Common Terminology Criteria for Adverse Events
ECG	Electrocardiogram
eCRF	Electronic case report form
IBS	Irritable bowel syndrome
IBS-GS	Irritable Bowel Syndrome Global Scale
IBS-D	Irritable bowel syndrome-diarrhea
IBS SS	Irritable Bowel Syndrome Severity Score
ITT	Intention-To-Treat Analysis Set
IWRS	Interactive web response system
kg	Kilograms
m	Meters
MedDRA	Medical Dictionary for Medical Affairs
mg	Milligrams
mm	Millimeters
mmHg	Millimeters of mercury
msec	Milliseconds
NTEAE	Non-treatment emergent adverse event
PD	Progressive disease or Pharmacodynamics
PK	Pharmacokinetics
PPS	Per- Protocol Analysis Set
PT	Preferred Term
Q1	25^{th} Percentile/First quartile
Q3	75^{th} Percentile/Third quartile
SAE	Serious adverse events
SAF	Safety
SAP	Statistical analysis plan
SAS®	Statistical Analysis System
SD	Standard deviation
SI	International System of Units

SMC	Safety monitoring committee
SOC	System Organ Class
TEAE	Treatment-emergent adverse events
T_{max}	Time to maximum observed plasma concentration
WAP	Worst abdominal pain
WHO	World Health Organization

1. INTRODUCTION

The statistical analysis plan (SAP) details the planned analysis required to satisfy the Clinical Study Report of study number BOS-589-201: "A Phase 2a, Randomized, Double-blind, Placebo-controlled, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of BOS-589 in the Treatment of Patients with Diarrhea-predominant Irritable Bowel Syndrome (IBS-D)". The content of this SAP is based on the protocol amendment 2 (version 3.0) dated 12 Aug 2019.

Mock shells for tables, listings, and figures are in a separate document. Analyses for pharmacokinetics (PK) and pharmacodynamics (PD) will be governed by a separate SAP.

2. STUDY OBJECTIVES AND ENDPOINTS

OBJECTIVES	ENDPOINTS
Primary	
• To evaluate in patients with IBS-D the abdominal pain response to BOS-589 after 4 weeks of treatment, relative to Placebo.	• 24-hour worst abdominal pain scores (WAP) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the overall safety and tolerability of BOS-589 in the treatment of IBS-D during 4 weeks of treatment, relative to Placebo.	• Incidence of adverse events (AEs) serious adverse events (SAEs), discontinuations because of AEs, and any treatment-related severe AEs.
Secondary	
• To evaluate the treatment effect of BOS-589 on defecation after 4 weeks, relative to Placebo.	• Change in stool consistency, measured by the daily Bristol Stool Form Score (BSFS) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). • Change in stool frequency, measured by the total number of spontaneous bowel movements in 24 hours at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the treatment effect of BOS-589 on IBS-related signs and symptoms.	• Change in the IBS Severity Score (IBS-SS) at Day 29 compared to baseline. • Change in the IBS Global Scale (IBS-GS) at Day 29 compared to baseline (averaged over the week prior to each respective timepoint).
• To evaluate the steady-state PK of BOS-589.	• Maximum observed plasma concentration (C_{max}); time to C_{max} (T_{max}); minimum plasma concentration (C_{min}); area under the concentration versus time curve (AUC) from time zero to 4 hours postdose (AUC_{0-4}); AUC from time zero to the last quantifiable concentration (AUC_{0-t}).

Additional Objectives	Additional Endpoints
Exploratory (not included in protocol amendment 2 [version 3.0] dated 12 Aug 2019)	
• To explore the treatment effect of BOS-589 in terms of weekly responders after 4 weeks of treatment, relative to Placebo.	• Number of patients who are weekly responders, defined by both an abdominal pain intensity weekly responder (decrease in the weekly

Additional Objectives	Additional Endpoints
	average of the WAP in the past 24 hours score of at least 30%compared with baseline averaged over the week prior) and a stool consistency weekly responder (at least 50%reduction in the number of days per week with at least one stool that has a consistency of Type 6 or 7 compared with baseline averaged over the week prior), at each scheduled visit.

3. STUDY DESIGN

3.1. Study Design and Population

This study is a phase 2a, randomized, double-blind, Placebo-controlled, multicenter trial designed to provide the first proof-of-principle efficacy of BOS-589 in IBS-D patients, and to inform dose selection for subsequent development. The study will consist of a pre-treatment phase, a 4-week double-blind treatment phase, and a 2-week post-treatment follow-up period.

Pre-Treatment Phase: During the pre-treatment phase, patients will be evaluated for up to 5 weeks to assess eligibility. The pre-treatment phase will consist of initial screening assessments and a Run-in period.

After the initial screening assessments have been performed, eligible patients will enter a Run-in period of up to 3 weeks. During the Run-in period, the patients will complete an electronic diary to collect daily information related to their IBS-D symptoms, bowel function, and rescue medicine use, to confirm disease activity and diary compliance. Patients will also be requested to discontinue any prohibited medications during this phase of the study.

Upon completion of the Run-in period, patients will return to the study site to confirm eligibility for randomization into a 4-week double-blinded Treatment Phase.

Treatment Phase: Eligible patients will be randomized (1:1:1) into the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for 200 mg BOS-589 twice daily (bid) orally;
- Cohort 2 (Low Dose): 2 x 25 mg tablets for 50 mg BOS-589 bid orally;
- Cohort 3 (Placebo): 2 x visually matching Placebo oral tablets bid.

During the 4 weeks of double-blind treatment, patients will continue to record their daily IBS-D symptoms, bowel function, and rescue medicine use in the electronic diary, as described in Section 8.3 of the protocol.

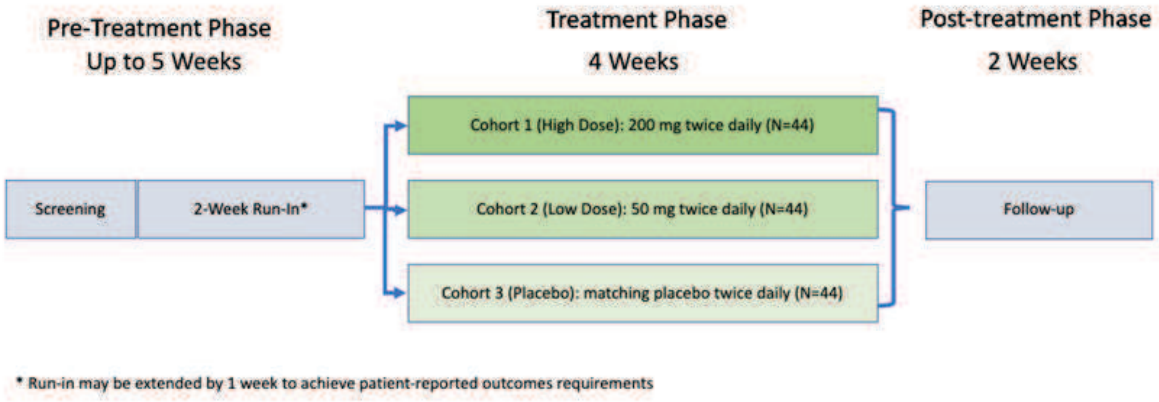
Post-treatment Phase: A 2-week post-treatment follow-up visit will occur for patients who complete the Treatment Phase. During the 2 weeks of follow-up, patients should continue to record their daily IBS-D symptoms, bowel function, and rescue medicine use in the electronic diary.

Patients who prematurely discontinue treatment should return to the study center to complete the early termination assessments as soon as possible after stopping the treatment.

Data analyses will occur after all patients in the trial have completed the last visit or procedure shown in the Schedule of Activities, Section 1.3 of the protocol.

Post-study Access to Therapy: No post-treatment access to therapy will be provided to patients randomized in the study.

Figure 1. Study Schematic



Further information on the study design, including eligibility criteria and a schedule of assessments, are detailed in the study protocol.

3.2. Changes to the Analyses from the Protocol

One exploratory efficacy endpoint was added in addition to the endpoints defined in the protocol.

- Summaries of weekly responders (defined in section 7.4) will be summarized as an exploratory efficacy endpoint.

The wording of how randomization file is handled is clarified in the SAP by saying the file will be held by an independent unblinded Biostatistician, not the study statistician.

3.3. Randomization and Blinding

This is a randomized, double-blinded, Placebo-controlled study. Patients should be randomized as close as possible to the time of the planned first dose of study drug.

Randomization will be conducted via an interactive web response system (IWRS) in a 1:1:1 ratio. Eligible patients will be randomized (1:1:1) into the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for 200 mg BOS-589 twice daily (bid) orally;
- Cohort 2 (Low Dose): 2 x 25 mg tablets for 50 mg BOS-589 bid orally;
- Cohort 3 (Placebo): 2 x visually matching Placebo oral tablets bid.

Patients will be assigned to the cohorts using a permuted block randomization stratified by site (44 patients per cohort). The randomization schedule will be generated by an independent unblinded Biostatistician and will be transferred to the IWRS team for loading into the system. PPD blinded biostatistics and blinded programming team will remain blinded. The randomization file will be held by the independent unblinded Biostatistician until the end of the trial and when the database has been locked.

The sponsor, site staff, and patients will be blinded to the treatment assignment. BOS-589 and Placebo will be similar in appearance, and provided in white, opaque, high-density polyethylene bottles, each with a child-resistant closure. Each bottle will be labeled as per country requirement.

Only in a medical emergency will a patient's treatment assignment be unblinded, and this process will be performed and documented in the IWRS. Every effort should be made to contact the Medical Monitor to discuss unblinding prior to breaking the blind.

3.4. Sample Size Considerations

This is a phase 2a study in which a hierarchical hypothesis test will be employed for the primary efficacy endpoint. The primary endpoint is defined as a change in the 24-hour WAP score at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). Mean change from baseline will be measured as a continuous variable and compared between cohorts in a hierarchical testing order using a t-test of equal variances and a two-tailed alpha of 0.05.

The sample size was based on the first hypothesis test in the hierarchical plan, with approximately 80% power, two-sided alpha of 0.05, and a standard deviation (SD) of 3 points on the numeric rating scale to detect a 1.6 minimum change between cohorts (N = 120 patients). To account for attrition and a potential higher Placebo rate, a total of 132 patients will be randomized to one of 3 cohorts in a 1:1:1 ratio (44 patients per cohort).

3.5. Safety Monitoring Committee

There is no safety monitoring committee reviewing data for this study.

3.6. Interim Analysis

No formal efficacy interim analysis is planned for the purposes of statistically comparing the data prior to the completion of the study.

3.7. Timing of Analyses

Formal hypothesis testing will occur at the end of the study and post database lock. There are no planned Data, Safety, or Adjudication Committees for this study and no planned statistical interim analyses.

4. DATA ANALYSIS CONSIDERATIONS

Unless noted otherwise, the following statistical considerations will be applied:

- All analyses will be conducted based on Statistical Analysis System (SAS®) software version 9.4 or higher.
- Data for all sites/centers will be pooled for analysis.
- All data in the database except birth date will be presented in by-patient data listings.
- Listings will be sorted by cohort, patient identification number (concatenated site and patient number), and assessment date (and time, if available). If assessment date is missing, chronological visit will be used for sorting.
- Continuous data will be summarized by cohort (or last assigned dose level, as applicable) based on n, mean, median, standard deviation (SD), 25th percentile/first quartile (Q1), 75th percentile/third quartile (Q3), minimum value, and maximum value.
- Categorical data will be summarized by cohort using n and percentage based on the number of non-missing values.
    - The number of missing values will be presented as a separate category with no percentage, but only if one or more patients are missing data.
    - Counts of zero will be presented without percentages.
- Statistics will be presented in the summary tables based on the following:
    - Mean, Median, Q1, and Q3: one additional decimal place to that reported for Minimum and Maximum
    - SD: two additional decimal places than the Minimum and Maximum

    - Percentages: reported to one decimal place
    - P-values will be reported to four decimal places. If the value is below 0.0001 it will be noted as < 0.0001; if the value above 0.9999 it will be noted as > 0.9999.
- The number and percentage will be presented in the form of XX (XX%). Percentage values of less than 1% will be presented as XX (<1%).
- Statistical inference will be based on a 5% significance level (i.e., 95% confidence intervals [CIs] will be produced). CIs will be produced only in cases with sufficient sample size (i.e., n => 5).
- Change from baseline = value (post-baseline visit) – value (baseline or last visit prior to treatment).
- All data up to the time of study completion/withdrawal from the study will be included in the analysis, regardless of duration of treatment.
- Numbering for data displays will be based on the International Council for Harmonization E3.
- The protocol uses the terms cohort, treatment group, and treatment arm interchangeably. For consistency in this document, the term cohort will be used.

4.1. Stratification and Covariates

There are no formal plans for analysis stratification or use of covariates.

4.2. Evaluation of Subgroups

No subgroup analyses are powered for the study. However, analyses for the primary and secondary endpoints will be analyzed by sex, and fecal calprotectin < 50 mcg/g at screening, and possibly other subgroups based on emerging data.

4.3. Multiple Comparisons and Multiplicity

A hierarchical hypothesis test will be employed for the primary efficacy endpoint, in the following pre-defined testing order using a t-test of equal variances and a two-tailed alpha of 0.05:

Hypothesis 1:

Active Treatment (Cohort 1 + Cohort 2) versus Placebo (Cohort 3);

If statistically significant at P < 0.05, then proceed to Hypothesis 2.

Hypothesis 2:

Cohort 1 (High Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05, then proceed to Hypothesis 3.

Hypothesis 3:

Cohort 2 (Low Dose) versus Cohort 3 (Placebo);

If statistically significant at $P < 0.05$ then, proceed to Hypothesis 4.

Hypothesis 4:

Cohort 1 (High Dose) versus Cohort 2 (Low Dose).

5. GENERAL DATA HANDLING CONVENTIONS

5.1. Assigned and Actual Treatment

Patients will be randomized via IWRS in a 1:1:1 ratio to receive BOS-589 200 mg, BOS-589 50 mg, or Placebo. Patients will be assigned to the cohorts using a permuted block randomization stratified by site (44 patients per group).

The assigned treatment is determined by a randomization schedule. The randomization schedule will be generated and will be transferred to the IWRS team for loading into the system. The actual treatment given to individual patient is the actual treatment.

All safety analyses will be based on the actual treatment. Study patient data and all efficacy analyses will be based on the assigned treatment unless specified otherwise.

5.2. Reference Dates

Reference dates are defined as follows:

- Screening date is defined as the electronic case report form (eCRF) provided visit date on which a patient was screened for trial entry.
- Informed consent date is defined as the eCRF provided date of informed consent from the informed consent page.
- Run-in period start date is defined as earliest date of eCRF provided visit date from the Run-in period page.
- Run-in period end date is defined as eCRF provided date of randomization -1 from the randomization page.
- Treatment start date is defined as the date of first dose of any study drug.
- Treatment end date is defined as the date of last dose of any study drug.
- Age will come directly from the eCRF which is integrated from IWRS based on the informed consent date.
- Safety data, such as AEs and laboratory assessments, will use the treatment start date as a reference date.
- Efficacy data will use the treatment start date as a reference date.

- Study day will be based on treatment start date as a reference date. Study Day 1 will be the day patients start the first dose of study drug.

5.3. Study Day and Duration Variables

Reference date calculations will generally be defined as the following, assuming non-missing dates:

- date of interest – reference date + 1 when the date of interest >= reference date;
- otherwise, date of interest – reference date.

If either date is missing, reference date calculations will not be performed. Date imputation will be performed as identified in Section 5.6.

Study day will either have a negative value if collected before dosing or a positive value if collected on or after the day of drug dosing; there will be no study day zero.

Duration of time is dependent on reference dates and will be calculated in a manner similar to that of the reference date calculation. Duration on study is defined as the end of study date – informed consent date + 1. Duration of treatment is defined as treatment end date – treatment start date + 1, where treatment end date is the date of last dose of study drug. Patients still receiving ongoing treatment or participating in study follow up at the time of analysis will use imputed treatment end and end of study dates as described in Section 5.6.

Durations reported in days will be calculated as measure date – reference date + 1. When reporting in months the result will be divided by 30.4375; for reporting in weeks it will be divided by 7; and for reporting in years it will be divided by 365.25.

5.4. Study Time Periods

Where applicable, data reporting will be classified by the following study periods for analysis:

- Pre-treatment is defined as the period prior to a patient's treatment start date.
- On-treatment is defined as the period between a patient's treatment start date and treatment end date. Treatment end date is defined as the date of last dose of study drug.
- Post-treatment is defined as the period of time following the on-treatment period.

5.5. Baseline and Post-Baseline Changes

Unless stated otherwise, baseline and post-baseline change values will be based on the following:

- Baseline will be based on the last non-missing value (including unscheduled assessments) collected prior to or on the treatment start date/time for measures collected prior to the treatment.
- Post-baseline values will be those collected after the date/time of first dose of study drug.

- Change from baseline is defined as: value – baseline value.
- Percentage change from baseline is defined as: (value – baseline value)/baseline value x 100%.
- Maximum change will be based on the change from baseline value to the maximum value for all post-baseline assessments collected during the on-treatment period, scheduled or unscheduled, using the formulas above.
- Minimum change will be based on the change from baseline value to the minimum value for all post-baseline assessments collected during the on-treatment period, scheduled or unscheduled, using the formulas above.

5.6. Imputation of Partial Dates

Appendix 1 details partial date conventions for AEs and medications.

Treatment End Date

Missing treatment end dates will not be imputed at the end of the study. However, due to ongoing reporting needs, treatment end date will be imputed as earliest of the data cutoff date, date of death, or last treatment date recorded on the exposure eCRF.

End of Study Date

Missing end of study dates will not necessarily be imputed at the end of the study. However, due to ongoing reporting needs, end of study dates will be imputed as the earliest of the data cutoff date, date of death, or last date recorded on the eCRF.

5.7. Multiple Assessments and Visit Windows

If there are multiple assessments at a scheduled visit, then the last assessment will be used in the analysis.

Nominal visits (e.g., those identified by the study eCRF) will be the basis of summarization and statistical analysis; visit windows for each scheduled visit can be found in the Schedule of Activities from the protocol. Unscheduled data may be included in summaries of baseline and maximum/minimum change values; summaries of specific abnormalities any time post-baseline; and patient data listings.

5.8. Lost to Follow Up or Lapse of Adequate Assessments

A patient will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site staff. Missing evaluations will not be imputed.

5.9. Cohort Display

Cohorts will be displayed with the following columns. See the mock shells for additional details and a visual representation of the cohort display.

- 200 mg BOS-589 bid
- 50 mg BOS-589 bid
- Active Treatment (Combined 200 mg and 50 mg Cohorts)
- Placebo bid
- Overall (not needed for efficacy outputs)

For primary efficacy analyses testing, cohort significant p-values will be displayed in the hierarchical order below. Following the first p-value, each subsequent p-value will only be displayed if the prior p-value is < 0.05.

- Active Treatment (Combined 200 mg and 50 mg Cohorts) vs. Placebo
- 200 mg BOS-589 bid vs. Placebo
- 50 mg BOS-589 bid vs. Placebo
- 200 mg BOS-589 bid vs. 50 mg BOS-589 bid

5.10. Missing Data

Missing data imputations for efficacy endpoints are described in Section 7; AE and concomitant medication, ongoing end of treatment, and ongoing end of study are described in Section 5.6. Otherwise, missing data will not be imputed.

6. STUDY PATIENT DATA

6.1. Analysis Sets

- The Screened Analysis Set will include all patients who have signed an informed consent document.
- The Intention-to-Treat (ITT) Analysis Set will include all randomized patients.
- The Per Protocol (PP) Analysis Set will include all patients in the ITT Analysis Set who received at least 1 dose of study drug and did not have any major protocol deviations.
- The Safety (SAF) Analysis Set will include all patients who received at least 1 dose of study drug.

All safety analyses will be performed in the SAF Analysis Set by the actual cohorts. All efficacy analyses will be performed in the ITT and PP Analysis Sets by the randomized cohorts. Summaries of the Screened Analysis Set will be performed by the randomized cohorts.

The criteria for inclusion in the PP Analysis Set will be finalized and documented before database is locked and study unblinding. Additional information on inclusion into this population can be found in Section 6.3.

6.2. Patient Disposition

A summary of patient enrollment by site will be presented by cohort and overall for the Screened Set. The overall patient enrollment status (screened [i.e., consented], screen failures, randomized) will be provided. Patients who sign the ICF and are randomized but do not receive the study drug may be replaced. Once randomized, patients who are withdrawn from therapy before completing 28 days of dosing or are discontinued from the study for any reason will not be replaced.

Disposition data will be summarized for the Screened Set. The number of patients in each analysis set (ITT, SAF, PPS), completed treatment, with ongoing treatment, completed the study, and still ongoing in the study, will be summarized. The number of patients who discontinued study drug including reasons, the duration on study drug, and duration on study will be summarized. The number of patients who discontinued the study including reasons, the duration on study drug, and duration on study will also be summarized. Data will be presented by cohort and overall by assigned treatment.

A by-patient listing of patient enrollment data as well as a by-patient listing of patient disposition data including reason for treatment and study discontinuation will be presented for the screened patients.

6.3. Protocol Deviations

Protocol deviations will be recorded in the clinical trial management system.

Protocol deviations will be identified and classified as major protocol deviations prior to database lock and final unblinding. Major deviations are those deviations that significantly impact efficacy and result in the removal of patients from the PP Analysis Set. Major protocol deviations may include but are not limited to:

- Violation of Inclusion/Exclusion Criteria
- Study drug compliance $\leq 80\%$ or $\geq 105\%$
- Use of prohibited therapies
- Incorrect treatment

Major protocol deviations will be summarized by cohort and deviation category. Protocol deviations will be listed for the ITT Analysis Set by assigned treatment.

6.4. Demographic, Baseline Characteristics, and Disease History

Patient demographics and baseline characteristics will be summarized by cohort and overall for the ITT Analysis Set by assigned treatment. These will include age, sex (Male, Female), gender (Male, Female, Transgender Female to Male, Transgender Male to Female, Not willing to respond, Other), childbearing potential for females (Yes, No), ethnicity (Hispanic or Latino, Not Hispanic or Latino), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander), height (cm), weight (kg), and body mass index (BMI; kg/m^2). BMI will be calculated as (weight in kg) / (height in m)^2.

Demographics and baseline characteristics will be listed for the ITT Analysis Set.

Disease history will be summarized by cohort and overall for the ITT Analysis Set by assigned treatment. These will include frequency of pain in the last 3 months, how often this pain happens close in time to a bowel movement, how often when the pain happens are stools softer or harder than usual, how often when the pain happens are stools more frequent or less frequent than usual, how often the pain gets worse with menstrual bleeding for females, how often the pain gets worse after a meal, how often the pain limit usual activities, if the pain is continuous or not, if the pain has lasted for over 6 months or not, and abnormal stool condition. These are all collected from eCRF.

Disease history data will be listed for the ITT Analysis Set in by-patient listings.

6.5. Medical History

Medical history will be summarized by Medical Dictionary for Medical Affairs (MedDRA) system organ class (SOC) and preferred term (PT). Sort order will be by descending overall incidence of SOC followed by descending incidence of overall PT. Medical history will be summarized and presented by cohort and overall for the SAF by actual treatment.

Medical history will be presented in data listings for the SAF Analysis Set. Medical history coded terms will be provided, including the SOC and PT per MedDRA. The data management plan specifies the version of MedDRA used for medical history coding.

6.6. Prior and Concomitant Medication

The incidence of medication use will be summarized by the World Health Organization (WHO) Drug Dictionary Anatomic Therapeutic Chemical (ATC) Level 2 classification (i.e., therapeutic main group) and preferred name. A patient will be counted only once at each level of reporting. Prior medications are those which have been identified to have been discontinued prior to the first dose of study drug (i.e., taken exclusively during the pre-treatment period). Concomitant medications are those which have been identified to have been taken at any point during the on-treatment period, including medications which started prior to first dose of study drug but are ongoing at first dose. The data management plan specifies the version of WHO Drug used.

Partial dates will be imputed according to Appendix 1 for the determination of prior, concomitant, and subsequent medications.

Concomitant medication use will be summarized and presented by cohort and overall for the SAF Analysis Set by actual treatment. Prior medications will be summarized in a similar manner.

During the treatment period, patients may also take protocol-permitted loperamide rescue medication for the acute treatment of uncontrolled diarrhea. Rescue medications will be summarized in a similar manner as the concomitant medication.

All prior and concomitant data will be provided together in a by-patient listing including the verbatim and preferred drug name and WHO ATC Class (ATC Level 2) for the SAF. Prior and subsequent medications will be flagged by unique identifiers in the listing. Loperamide rescue medication use will also be flagged in the listing.

6.7. Rescue Medication

Rescue medication usage (loperamide) will be summarized by cohort and overall for all patients in the SAF Analysis Set. The time periods to be summarized will include the run-in period prior to Baseline/Day 1, the treatment period (Baseline/ Day1 through Day 29), and the post-treatment period (after Day 29 through Day 43). Categorical summaries will include the incidence of patients with rescue medication use at any time during the associated time period, as well as, the number of total doses of rescue medication used (0 doses, 1 dose, 2 doses, 3 doses, 4 doses, 5 doses, 6 doses, 7 doses, 8 doses, 9 doses, 10 or more doses) during the associated time period. The total number of doses of rescue medication also will be summarized as a continuous measure during the associated time period.

6.8. Study Drug Exposure and Compliance

Exposure to study drug will be summarized by cohort and overall for all patients in the SAF Analysis Set.

Patients will be randomized to receive for a total of 4 weeks BOS-589 at 1 of 2 dose levels or Placebo bid, as described in the following cohorts:

- Cohort 1 (High Dose): 2 x 100 mg tablets for the 200 mg BOS-589 bid orally
- Cohort 2 (Low Dose): 2 x 25 mg tablets for the 50 mg BOS-589 bid orally
- Cohort 3 (Placebo): 2 x visually matched Placebo tablets bid orally

The duration of exposure to study drug is calculated as date of last dose of study drug taken – date of first dose of study drug taken +1.

The planned total number of tablets during exposure is defined as duration of exposure x 4.

The actual total number of tablets taken during exposure is calculated as the total number of tablets given – the total number of tables returned, both collected in the eCRF.

The average actual daily number of tablets taken is calculated as the actual total number of tablets taken during exposure / duration of exposure to study drug.

The drug compliance is calculated as the actual total number of tablets taken during exposure / planned total number of tablets taken during exposure x 100%.

The duration of exposure to study drug, planned and actual total number of tablets during exposure, average actual daily number of tablets taken, and drug compliance will be summarized. Drug compliance categories of >=105%, >100% to <105%, >90% to <=100%, >80% to <=90%, >70% to <=80%, and <=70% will also be presented.

A by-patient listing of exposure and compliance measures will be produced for SAF Analysis Set.

7. EFFICACY

Efficacy analyses will be conducted using the ITT and PP Analysis Sets.

All available efficacy data will be provided in by-patient data listings on all ITT patients.

The efficacy objectives of the study are to evaluate the effect of BOS-589, relative to Placebo after 4 weeks of therapy, on pain, defecation, and key IBS-D related signs and symptoms.

During the 4 weeks of the double-blind treatment phase, patients will be required to access their electronic diary every day, preferably at the same time each day, to record IBS-D symptom data and information related to their bowel function and rescue medication use.

For all efficacy endpoints, the Day 29 change from baseline (with scores averaged over the week prior to each respective timepoint) is the main timepoint of interest. However, unless stated otherwise, data will be presented for each scheduled timepoint (Baseline/Day 1, Day 8, Day 15, Day 22, Day 29, and Day 43). The baseline score will be the average of non-missing scores reported for the last 7 days prior to and inclusive of the date where eligibility and run-in criteria pertaining to IBS-D symptom and bowel function have been met. The Day 29 score will be the average of non-missing scores reported from Day 22 to Day 28. The Day 8, Day 15, Day 22, and Day 43 timepoints are calculated in a similar manner. If there are missing scores on any day, the weekly average will be taken from all the non-missing values. No imputation of the missing values will be made.

7.1. Primary Efficacy Endpoint and Analyses

7.1.1. Worst Abdominal Pain

The primary efficacy endpoint will be defined as a change in the 24-hour WAP score at Day 29 compared to baseline (averaged over the week prior to each respective timepoint). The 24-hour WAP score will be a patient-reported score of their WAP in the past 24 hours on a 0 to 10 scale, where 0 corresponds to no pain and 10 corresponds to worst imaginable pain. Patients will be asked to rate their WAP in the past 24 hours every day for the entire Run-in period and the 4-week double-blind treatment period.

This is a proof-of-principle study in which a hierarchical hypothesis test will be employed for the primary efficacy endpoint. The mean change in 24-hour WAP at Day 29 from baseline (averaged over the week prior to each respective timepoint) will be measured as a continuous variable and compared between cohorts in the following testing order using a t-test of equal variances and a two-tailed alpha of 0.05:

Hypothesis 1:

Active Treatment (Cohort 1 + Cohort 2) versus Placebo (Cohort 3);

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

The null hypothesis will be that there is no difference in the mean change in 24-hour WAP score at Day 29 from Baseline (averaged over the week prior to each respective timepoint) between active treatment and Placebo. The alternative hypothesis will be that there is a difference in the mean change in 24-hour WAP score at Day 29 from Baseline (averaged over the week prior to each respective timepoint) between active treatment and Placebo. The test statistic is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

$\bar{x}_1$ and $\bar{x}_2$ are the sample means of change at Day 29 from Baseline in 24-hour WAP score (averaged over the week prior to each respective timepoint), s^2 is the pooled sample variance assuming equal variance, n_1 and n_2 are the sample sizes, and the degree of freedom is $n_1 + n_2 - 2$.

If statistically significant at $P < 0.05$, then proceed to Hypothesis 2.

Hypothesis 2:

Cohort 1 (High Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05, then proceed to Hypothesis 3.

Hypothesis 3:

Cohort 2 (Low Dose) versus Cohort 3 (Placebo);

If statistically significant at P < 0.05 then, proceed to Hypothesis 4.

Hypothesis 4:

Cohort 1 (High Dose) versus Cohort 2 (Low Dose).

Methodology for hypotheses 2 to 4will be similar to that described for hypothesis 1.

The patient-reported 24-hour WAP scores and change from baseline scores (averaged over the week prior to each respective timepoint) will be summarized descriptively by cohort for each scheduled visit. Box plots of the 24-hour WAP scores (averaged over the week prior for each respective timepoint) will also be presented by cohort at each scheduled visit. Line plots of 24-hour WAP change from baseline scores (averaged over the week prior for each respective timepoint) will also be presented by cohort at each scheduled visit.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of worst abdominal pain will be repeated as a sensitivity analysis on this subgroup.

7.2. Secondary Efficacy Endpoints and Analyses

7.2.1. Stool Consistency

Patients will be asked to record daily stool consistency according to the BSFS most representative stool consistency in the past 24 hours and the worst stool consistency (defined as the loosest stool with the highest BSFS score) in the past 24 hours. The patient-reported BSFS consistency score, based on a 1 to 7 scale where 1 corresponds to a hard stool and 7 corresponds to watery diarrhea (CCI [redacted]), can be found in Appendix 2.

The patient-reported stool consistency scores and change from baseline scores (averaged over the week prior to each respective timepoint) will be summarized descriptively by cohort for each scheduled timepoint. Pooled active treatment data will be compared to Placebo for the analyses, and by cohort, using a t-test of equal variances and a two-tailed alpha of 0.05 specified in section 7.1.1 for reference. Box plots of the BSFS scores (averaged over the week prior to each respective timepoint) will also be presented by cohort for each scheduled visit. Line plots of the BSFS change from baseline values (averaged over the week prior to each respective timepoint) will also be presented by cohort for each scheduled visit. Analyses for the most representative stool consistency in the past 24 hours and worst (loosest) stool consistency in the past 24 hours will both be presented.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of BSFS most representative consistency and BSFS worst consistency will be repeated as sensitivity analyses on this subgroup as described in section 7.1.1.

7.2.2. Stool Frequency

Patients will be asked to record stool frequency based on the total number of spontaneous bowel movements in the past 24 hours. Bowel movement frequency and change from baseline (averaged over the week prior to each respective timepoint) will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of stool frequency will be repeated as a sensitivity analysis on this subgroup as described in section 7.1.1.

7.2.3. Irritable Bowel Syndrome Severity Score

IBS-SS will be collected at baseline and Day 29 only. Patients will be asked to complete 5 questions regarding the severity of their IBS. Each of the 5 questions generate a maximum score of 100, leading to a total possible score of 500. IBS-SS will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1 (except only for scheduled visits of Baseline/Day 1, Day 29, and Day 43). Summaries performed by visit will include responses to the individual questions, as well as the total scores.

7.2.4. Irritable Bowel Syndrome Global Score

Patients will be asked to record daily their overall IBS-GS in the prior 24 hours. The patient-reported daily IBS-GS is based on a 0 to 4 scale where:

- 0 corresponds to no symptoms;
- 1 corresponds to mild symptoms;
- 2 corresponds to moderate symptoms;
- 3 corresponds to severe symptoms;
- 4 corresponds to very severe symptoms.

IBS-GS will be analyzed and summarized using the same method as the secondary endpoint described in Section 7.2.1 (except only for scheduled visits of Baseline/Day 1 and Day 29).

7.3. Exploratory Efficacy Endpoints and Analyses

7.3.1. Weekly Responder

Weekly responders will be summarized as an exploratory efficacy endpoint.

A patient is categorized as a weekly responder if the patient is a weekly responder in both pain intensity and stool consistency (page 8 of FDA guidance document on IBS).

- An Abdominal Pain Intensity Weekly Responder is defined as a patient who experiences a decrease in the weekly average of the WAP in the past 24 hours score of at least 30 percent compared with weekly average at baseline, i.e., (weekly average of baseline – weekly average of a post-baseline visit) / weekly average of baseline >= 0.3
- A Stool Consistency Weekly Responder is defined as a patient who experiences a 50 percent or greater reduction in the number of days per week with at least one stool that has a consistency of Type 6 or 7 compared with weekly average at baseline, i.e., (weekly average of the number of Type 6 or 7 BSFS stool of baseline – weekly average of the number of Type 6 or 7 BSFS stool of a post-baseline visit) / weekly average of the number of Type 6 or 7 BSFS stool of baseline >= 0.5

The number and percentage of patients who are abdominal pain intensity weekly responders, stool consistency weekly responder, and weekly responders in both will be summarized by cohort at each scheduled visit. A chi-square test will be used, as a reference, to test whether there is a difference by between each active dose level and Placebo, both active dose levels and Placebo, and between the two active dose levels. Fisher's exact test will be used if the expected values in any of the cells of the contingency table are below 5.

To evaluate the treatment effects in the subset of patients with fecal calprotectin under 50mcg/g at screening, this analysis of weekly responders will be repeated as a sensitivity analysis on this subgroup.

8. PHARMACOKINETICS / PHARMACODYNAMICS

PK/PD analyses are covered under a separate SAP.

9. SAFETY

Unless otherwise specified, all safety analysis reporting will be based on the SAF Analysis Dataset. For safety analyses, Day 29 represents measures taken at end of treatment/early termination while Day 43 represents measures taken at the end of study.

9.1. Adverse Events

AEs starting after the informed consent date until the Day 43/End of Study Follow-up visit will be recorded. AE assessments will include severity (mild, moderate, severe), relationship to study drug (related or not related), and seriousness. AEs with a missing severity will be classified as severe. AEs with a missing causality will be classified as related. AEs with missing seriousness will be classified as serious.

Any event reported on the CRF that occurs on or after the initiation of study drug is defined as a treatment-emergent adverse event (TEAE). Guidelines for categorizing an event as treatment-emergent based on start and stop dates and times are in Appendix 1. Additionally, an AE that is reported to have started on Day 1 without an associated onset time will be assumed to have occurred after the initiation of study drug. Hence, AEs occurring on Day 1 with no associated onset time will be assumed to be treatment emergent. SAEs associated with a protocol specified procedure and occurring after the time of consent but before administration of first dose of study drug will be defined as non-treatment emergent AEs (NTEAEs) and will be excluded from summary tables but will appear in listings flagged as NTEAE.

On a summary table, the number and percentage of patients in each of the following categories will be summarized by cohort following the cohort display in section 5.9:

- Any TEAE
- Any TEAEs related to study drug
- Any moderate or severe TEAEs
- Any moderate or severe TEAEs related to study drug
- Any treatment-emergent SAEs
- Any treatment-emergent SAEs related to study drug
- Any TEAEs causing discontinuation of study drug
- Any TEAEs with outcome of death
- Any TEAEs leading to withdrawal from the study

AEs will be coded to a SOC and PT using MedDRA (the version used will be defined in the study data management plan).

For summaries of TEAEs by SOC and PT, a patient will be counted once within an SOC, even if the patient experienced more than one TEAE within a specific SOC (likewise for PT). TEAEs will be ordered by descending frequency for overall SOC incidence followed by descending frequency of overall PT incidence. Summaries of TEAEs by SOC and PT will include the following:

- All TEAEs
- All TEAEs by severity
- All moderate or severe TEAEs
- All TEAEs related to study drug
- All treatment-emergent SAEs

- All treatment-emergent SAEs related to study drug
- All TEAEs causing discontinuation of study drug
- All TEAEs with outcome of death
- All TEAEs leading to withdrawal from the study

All AEs, including the verbatim and coded terms, will be listed for all enrolled patients. Additional separate listings will be provided for all AEs with outcome of death, all AEs causing discontinuation of study drug, all SAEs, and all AEs leading to withdrawal from the study.

9.2. Clinical Laboratory Evaluations

Laboratory tests identified in Appendix 1 in the protocol will be performed at times defined in the Schedule of Activities, Section 1.3 of the protocol. Clinical chemistry, hematology, and coagulation parameters will be reported based on the International System of Units (SI). All abnormal laboratory results will be evaluated by the Investigator as either clinically significant or not clinically significant.

Observed values and changes from baseline for hematology, clinical chemistry, urinalysis, and inflammatory biomarker laboratory evaluations will be summarized by cohort at each scheduled visit (Baseline/Day 1, Day 15, Day 29/Early Termination, and Day 43/End of Study) and most extreme change.

Box plots will be produced by visit and treatment for the recorded values of Erythrocyte Sedimentation Rate, C-reactive Protein, and Fecal Calprotectin. As well, line plots of the mean change from baseline of these three tests at each visit will be produced by treatment.

The number and percentage of patients experiencing treatment-emergent graded toxicities will be summarized by cohort and severity grade per the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.3. For each parameter, a patient will only be counted once using the worst-case value among all post-baseline timepoints. Laboratory toxicity will be summarized by incidence of CTCAE grade 3 and CTCAE grade 4 values for measures with a corresponding CTCAE grade. For measures without a corresponding CTCAE grade, incidence of 2.5 to < 5.0 X upper or lower limit of normal and ≥ 5.0 X upper or lower limit of normal cohort by cohort. Patients with missing values post-baseline will be excluded from the summary. The denominator for percentages will be the number of patients with at least one post-baseline assessment for the laboratory parameter being displayed. Separate summary counts will be prepared for parameters with bi-directional toxicity grading with "high/hyper" conditions depending on the maximum post-baseline value and the "low/hypo" conditions depending on the minimum post-baseline value.

All laboratory parameters will be provided in patient data listings for all enrolled patients.

9.3. Other Safety Evaluations

For electrocardiogram and vital signs results, Marked Abnormality will be defined in the following sections. Should a patient have a Markedly Abnormal result at baseline, only those

post-baseline results that are at least 1 grade more extreme than the baseline grade will be considered as a Marked Abnormality.

9.3.1. 12-Lead Electrocardiogram

12-lead electrocardiograms (ECGs) will be assessed as specified in the Schedule of Activities in Section 1.3 of the protocol. The following ECG parameters will be collected: PR interval (milliseconds [msec]), QRS interval (msec), QT interval (msec), QT interval corrected for heart rate (QTc) (msec), and Ventricular Rate (beats/min). ECG parameters and ECG evaluation will come directly from the database and will not be calculated during analysis.

Observed values and changes from baseline for ECG parameters will be summarized at each scheduled visit (Screening and Day 29/ Early Termination) by cohort and overall.

Observed and change from baseline values on Day 29 in QTc status categorized as markedly abnormal (defined in the table below) or not will be summarized.

QTc Measure	**Markedly Abnormal Criteria**
Observed	450–480 msec, inclusive [CTCAE grade 1] 481–500 msec, inclusive [CTCAE grade 2] > 500 msec [CTCAE grade 3 or higher]
Change from Baseline	31–60 msec, inclusive, increase from baseline >60 msec increase from baseline

If triplicate measures were taken at a single time point, the average will be summarized.
All ECG data will be presented in a by patient data listing for all enrolled patients. Measures meeting the markedly abnormal criteria will be flagged in the listing.

9.3.2. Physical Examinations

Physical examinations will be presented in patient data listings on all enrolled patients.

9.3.3. Vital Signs

Vital signs include: pulse rate (beats/min), respiratory rate (breaths/min); temperature (degrees Celsius [°C]); systolic and diastolic blood pressure (millimeters of mercury [mmHg]); height, weight (kg), and BMI (derived).

Observed values and changes from baseline for vital signs will be summarized at by cohort and overall for each scheduled visit (Baseline/Day 1, Day 8, Day 15, Day 22, Day 29, and Day 43) and most extreme change.

Summaries of pulse rate, temperature, systolic blood pressure, and diastolic blood pressure will be based on markedly abnormal criteria defined below:

Vital Sign	Markedly Abnormal Criteria
Pulse rate (beats/min)	< 60 beats/min > 100 beats/min
Temperature (°C)	≤ 35 °C ≥ 38 °C
Systolic blood pressure (mmHg)	120-139 mmHg, inclusive (CTCAE grade 1) 140–159 mmHg, inclusive (CTCAE grade 2) ≥ 160 mmHg (CTCAE Grade 3)
Diastolic blood pressure (mmHg)	80–89 mmHg, inclusive (CTCAE grade 1) 90–99 mmHg, inclusive (CTCAE grade 2) ≥ 100 mmHg (CTCAE grade 3)

Incidence of post-baseline markedly abnormal worst-case values will be presented. For pulse rate and temperature, both high and low values will be presented separately such that patients can be counted in both categories.

All vital signs data will be presented in patient data listings for all enrolled patients. Markedly abnormal vital sign values will be flagged as such in the vital signs listing.

10. REFERENCES

1. CCI [redacted]
2. FDA Guidance for Industry Irritable Bowel Syndrome — Clinical Evaluation of Drugs for Treatment 2012: https://www.fda.gov/media/78622/download

11. APPENDICES

APPENDIX1 Partial Date Conventions

ALGORITHM FOR TREATMENT EMERGENCE OF ADVERSE EVENTS:

START DATE	STOP DATE	ACTION
Known	Known/Partial /Missing	If start date < study drug start date, then not TEAE If start date >= study drug start date and < end of treatment + 15 in days , then TEAE
Partial, but known components show that it cannot be on or after study drug start date	Known/Partial /Missing	Not TEAE
Partial, could be on or after study drug start date	Known	If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Missing	Assumed TEAE
Missing	Known	If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, then not TEAE If stop date >= study drug start date, then TEAE
	Missing	Assumed TEAE

ALGORITHM FOR PRIOR / CONCOMITANT MEDICATIONS:

START DATE	STOP DATE	ACTION
Known	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Missing	If stop date is missing could never be assumed a prior medication If start date <= end of treatment + 15 in days, assign as concomitant
Partial	Known	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Partial	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown) and impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date and start date <= end of treatment + 15 in days, assign as concomitant
	Missing	Impute start date as earliest possible date (i.e., first day of month if day unknown or 1st January if day and month are unknown), then: If stop date is missing could never be assumed a prior medication If start date <= end of treatment + 15 in days, assign as concomitant
Missing	Known	If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant

START DATE	STOP DATE	ACTION
	Partial	Impute stop date as latest possible date (i.e., last day of month if day unknown or 31st December if day and month are unknown), then: If stop date < study drug start date, assign as prior If stop date >= study drug start date, assign as concomitant
	Missing	Assign as concomitant

APPENDIX2 The Bristol Stool Form Scale