Protocol 2017-01

A Phase 3 Prospective, Randomized, Double-blinded, Placebo-controlled Clinical Study to Evaluate the Efficacy and Safety of Rebiotix RBX2660 (microbiota suspension) for the Prevention of Recurrent *Clostridium difficile* Infection

NCT03244644

Statistical Analysis Plan

Version: 8.0
Date: 25 September 2020

Table of Contents

1 LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Definition
AIDS	Acquired Immunodeficiency Syndrome
AE	adverse event(s)
ATC	anatomical therapeutic chemical
CI	confidence interval
CDAD	*Clostridium difficile*-associated diarrhea
CDI	*Clostridium difficile* infection
Cdiff32	CDI-specific health-related quality of life survey instrument
CSR	Clinical Study Report
DSMB	Data Safety Monitoring Board
eCRF	electronic case report form
EAC	Endpoint Adjudication Committee
FDA	Food and Drug Administration
FT	fecal transplant
HIV	human immunodeficiency virus
HIPAA	Health Insurance Portability and Accountability Act of 1996
IBD	inflammatory bowel disease
IBS	irritable bowel syndrome
ID	study identification number
ITT	intent-to-treat
IP	investigational product (RBX2660 or placebo)
ICU	intensive care unit
IVIG	intravenous immunoglobulin
MedDRA	Medical Dictionary for Regulatory Activities
MITT	modified intent-to-treat
mL	milliliters
N	number of subjects
PP	per-protocol
PT	preferred term
RBX2660	investigational drug product being evaluated
SAE	serious adverse event
SAP	Statistical Analysis Plan
SD	standard deviation
SOC	system organ class
TEAE	treatment-emergent adverse event
US	United States
WHO	World Health Organization

2 INTRODUCTION

RBX2660 (microbiota suspension) is being studied for the prevention of recurrent *Clostridium difficile* infection (CDI) in adult subjects. The purpose of this study is to confirm the efficacy and safety of RBX2660 for the prevention of recurrent CDI in subjects who have had prior recurrent *Clostridium difficile* infection that was resolved with antibiotic treatment.

3 STUDY OBJECTIVES AND ENDPOINTS

3.1 Primary Objective

To confirm the efficacy of RBX2660 as compared to a Placebo in preventing recurrent episodes of CDI through 8 weeks.

3.2 Secondary Objective

The secondary objective of this study is:

To evaluate the sustained clinical response rate of RBX2660 as compared to Placebo after blinded treatment.

3.3 Other Objectives

1. To confirm the safety and tolerability of RBX2660.
2. To identify baseline characteristics predictive of efficacy outcomes.
3. To characterize the changes from baseline fecal microbial composition in subjects treated with RBX2660 as compared to Placebo.
4. To characterize the changes from baseline comorbidities in subjects treated with RBX2660 as compared to Placebo.
5. To evaluate health-related quality of life for CDI as measured by the Cdiff32 questionnaire.
6. To characterize the baseline severity of CDI in subjects with documented CDI recurrence.

7. Evaluate treatment success of RBX2660 in Placebo subjects who are documented study treatment failures then went on to receive RBX2660.
8. Assess the ability of more than one dose of RBX2660 to prevent CDI recurrence.
9. Assess the combined treatment success of all subjects receiving a single dose of RBX2660 during the study both to prevent recurrent CDI as well as prevent new CDI episodes.
10. Assess the clearance rate of vancomycin resistant enterococcus in subjects who are carriers at baseline.
11. Assess the clearance of *C. difficile* following enema treatment at 4 and 8-weeks and 3 and 6 months after blinded study treatment in subjects receiving RBX2660 and those receiving Placebo.

4 STUDY DESIGN

4.1 General Study Design

This is a prospective, multicenter, randomized, double-blinded, placebo-controlled Phase 3 study to confirm the efficacy and safety of RBX2660 for the prevention of recurrent CDI.

Up to 270 subjects may be randomized and treated in this study. Randomization will be at a 2:1 ratio (RBX2660:Placebo) which means approximately 180 subjects randomized to treatment with RBX2660 and 90 subjects randomized to Placebo. Consented subjects who meet all inclusion and none of the exclusion criteria will be randomized. Potential subjects who are consented but do not meet the inclusion and/or exclusion criteria will be considered screen failures and will not count towards total randomized. Subjects who are randomized but exit prior to the administration of the first blinded enema may be replaced with additional subjects which could cause overall number of subjects randomized to be greater than the 270 anticipated, however no more than 270 subjects will be treated with a blinded enema. Replacement subjects will also be randomized in order to ensure proper blinding. Subjects will be randomized to one of the two following study arms:

Placebo Arm – Group A

- Placebo dosing consists of one enema of placebo given in a blinded manner.

RBX2660 Arm – Group B

- Active treatment consists of one enema of RBX2660 given in a blinded manner.

The target population is adults (≥ 18 years old) with recurrent CDI who have had at least one recurrence after a primary episode (i.e., at least two episodes) and have completed at least one round of standard oral antibiotic therapy.

The definition of recurrent CDI for study entry is a documented diagnosis of:

1. CDI diarrhea that began within 8 weeks after completion of previous CDI treatment;
2. *and* at least one positive stool test for the presence of toxigenic *C. difficile*; which must be within 30 days prior to or on the date of enrollment (informed consent).

To be considered for enrollment, a subject must have medical record documentation of recurrent CDI per the above study definition that includes either:

a) at least one recurrence after a primary episode and has completed at least one round of standard-of-care oral antibiotic therapy or;

b) had at least two episodes of severe CDI resulting in hospitalization within the last year.

The definition of CDI diarrhea for use in enrolled subjects throughout the study includes:

1. The passage of three or more unformed/loose stools (i.e. Bristol Stool Scale type 6-7) in 24 or fewer consecutive hours for at least two consecutive days;
2. *and* a positive stool test for the presence of *C. difficile* toxin; documented at the time of the diarrhea.

Potential subjects are expected to already be taking or have just been prescribed antibiotics to control recurrent CDI symptoms at the time of enrollment per the investigator's standard of care. This means patients who have already completed their prescribed course of antibiotics to treat recurrent CDI, cannot be considered for enrollment unless they have another recurrence.

Once enrolled (informed consent obtained), in order to be randomized and treated;

1. Subject must meet all the inclusion and none of the exclusion criteria;
2. Antibiotics will have been administered for a minimum of 10 consecutive days prior to the washout period;
3. CDI symptoms must be under control leading into the antibiotic washout period.

Control of symptoms is defined as no longer meeting the symptomatic criteria for CDI diarrhea which is the passage of three or more unformed/loose stools (i.e. Bristol Stool Scale type 6-7) in 24 or fewer consecutive hours for at least two consecutive days while taking antibiotics. Thus, at minimum the Subject Diary (completed from time of enrollment until day of treatment) must show that symptoms are under control the two days prior to the washout period in order for the subject to be treated.

Once screening is completed and all inclusion and exclusion verified, the subject can be randomized. The assigned study treatment will be scheduled to consist of a single blinded study enema according to the randomization assignment. Treatment is to be completed as soon as possible, but no more than 14 calendar days following randomization. A minimum of 24hr to maximum of 72hr antibiotic washout period is required prior to administration of the assigned study treatment.

In-office study follow-up visits occur at weeks 1, 4 and 8 after completing the blinded study treatment. Telephone assessments for adverse events occur during weeks 2, 3 and 6 after the study enema and at months 3 and 6. Subjects are required to keep a detailed diary to assess for

solicited events from the date of enrollment to the 1-week follow-up visit after receiving the blinded study enema. The diary will be collected and reviewed at the baseline visit prior to blinded enema administration, and then collected and reviewed at the 1-week follow-up visit following enema administration.

Subjects who are deemed failures following the blinded treatment per the pre-specified treatment failure definition may elect to receive an unblinded RBX2660 enema. The unblinded enema is to be administered within 21 calendar days of failure determination. If a subject has a CDI episode on-study and the treatment plan is to pursue a treatment other than that specified in the protocol, the subject will be exited prior to receipt of the non-study treatment (ex. fecal microbiota transplant).

4.1.1 Treatment Success

Treatment success is defined as:

- The absence of CDI diarrhea for 8 weeks after completing a study treatment.

4.1.2 *Sustained Clinical Response*

Sustained clinical response is defined as:

- Treatment success of the presenting CDI recurrence and no new CDI episodes for greater than 8 weeks after completing a study treatment.

4.1.3 Treatment Failure

Treatment failure (CDI recurrence) is defined as:

- The presence of CDI diarrhea within 8 weeks of administration of a study enema; which includes a positive stool test for *C. difficile* toxin at the time of the diarrhea.

To ensure consistency across clinical sites for the determination of study failures, the central laboratory will perform the required C. DIFF QUIK CHEK COMPLETE® Test rapid enzyme

immunoassay for glutamate dehydrogenase (GDH) and toxins A and B. If the test is inconclusive, the central laboratory will conduct a second confirmatory test by PCR.

In the case of suspected treatment failure where there is an immediate concern for subject safety, including the potential for serious adverse events, a C. DIFF QUIK CHEK COMPLETE® Test will be permitted to be conducted at the site or other laboratory known to the site in addition to sending a sample to the central laboratory testing. If the C. DIFF QUIK CHEK COMPLETE® Test is not available, a testing algorithm that directly assesses the presence of *C. difficile* toxin for laboratory confirmation of recurrent CDI should be used. In the event, the test results from the central laboratory differ from that of a local laboratory, the results from the central laboratory will be used in determination of treatment failure. For subjects who are suspected of treatment failure, but for whom there is not a test result from the central laboratory for the suspected episode, a deviation will be entered to the database, and due documentation of the subject's symptoms and all available treatment and outcome data will be sent to the Endpoint Adjudication Committee (EAC) for blinded, independent adjudication of treatment outcome.

4.1.4 Determination of Treatment Failures

The site investigator first makes the initial determination of treatment success or failure based on the pre-defined study definitions. The site investigator's assessment will be provided to the EAC for blinded, independent adjudication of treatment outcome that will be utilized for study analysis and reporting purposes. This may be done on an on-going basis or a single meeting as deemed appropriate.

4.2 Blinding

Subjects, site personnel, the Data and Safety Monitoring Board (DSMB), EAC, and the Medical Monitor are initially blinded to the randomization assignment. Blinding will be maintained through the data cutoff for the primary efficacy analysis after which the blind will be lifted for the DSMB. The DSMB may be unblinded to individual assignments earlier as needed to adjudicate certain adverse events. Individual randomization assignments will not be revealed to

the subjects, site personnel, EAC or the Medical Monitor until the study is completed. Sites will be provided an emergency unblinding code if it becomes medically necessary to know the treatment group assigned for a particular subject.

Subjects will be randomized to receive either RBX2660 or placebo through an Interactive Response Technology (IRT) System prior to receiving the first assigned blinded study treatment. Randomization assignment is revealed to Rebiotix Operations Department personnel for product preparation and shipment. The Rebiotix Kit is shipped to the site in a sealed carton that is only opened by the IP administrator at the time of administration, out of view of the subject and other site personnel. While the IP administrator will not have access to the actual randomization assignment, it may be possible to guess through visualization at the time of enema preparation and administration. Additionally, subjects will be instructed to disregard any product characteristics that may become evident during or after administration such as the presence or absence of color or odor. To limit study bias, the IP administrator is not allowed to perform any of the subject study assessments.

4.3 Method of Assignment of Subjects to Treatment Arms

After enrollment and confirmation of inclusion/exclusion criteria compliance, randomization occurs via the study database with blinded treatment occurring within 14 calendar days of randomization. The randomization assignment is transmitted only to the Rebiotix Operations Department, who will ship the correct IP to the site for each treatment. Subjects are considered enrolled once they sign the informed consent form.

Trained and authorized study personnel will access the centralized randomization system to randomize the subject to one of the available treatment arms. The randomization schedule will be created using randomized blocks within four strata based on antibiotics used at screening (Vancomycin alone, Vancomycin in combination, Fidaxomicin or Other), and randomized subjects will be assigned one of two treatments (Group A or Group B) within the appropriate

stratum. The randomization will not be stratified by site, so each site will draw from the same set of blocks. A subject will be assigned the next available treatment allocation from the randomization list. The randomization code and the date and time of randomization will be captured in the clinical database. Randomization will be in a 2:1 ratio with approximately 2 subjects randomized to active treatment for every 1 subject randomized to placebo.

Enrolled subjects who are exited prior to randomization or are randomized but exited prior to attempting the blinded treatment may be re-enrolled and thus randomized under a different subject identification numbers (re-randomized). If this occurs, they will be included in the analysis using their second randomization assignment and subsequent assessments and will only be counted once towards the total number randomized.

4.4 Determination of Sample Size

To demonstrate a 69% success with one dose of the RBX2660 treatment group vs. 47% success in the Placebo control group, 240 subjects are required (power > 90%, nominal 2.5% type I error rate). Up to an additional 30 subjects will be enrolled to allow for a 10% loss-to-follow up rate, for a total of approximately 270 subjects. Considering the 2:1 randomization, this results in approximately 180 subjects in the RBX2660 arm and 90 subjects in the Placebo arm. Enrolled subjects who have signed consent and withdraw for any reason prior to administration of the first blinded enema will be replaced without counting toward the sample size cap. Replacement subjects will be randomized to ensure proper blinding.

4.5 Interim Sample Size Re-estimation

There are no plans for an interim sample size re-estimation being completed for this study.

4.6 Randomized Treatment Arms

Placebo Arm – Group A

- Placebo dosing consists of one enema of placebo given in a blinded manner.

RBX2660 Arm – Group B

- Active treatment consists of one enema of RBX2660 given in a blinded manner.

5 CHANGES IN THE CONDUCT OF THE STUDY OR PLANNED ANALYSES

5.1 Changes in the Conduct of the Study

There have been no changes to the conduct of the study or planned analysis.

6 BASELINE, EFFICACY AND SAFETY EVALUATIONS

6.1 Schedule of Events

Activity	Screening (Enrollment)	Baseline / Enema Administration (≤21 days from screening)	Follow-up Visits 1-, 4- and 8-Week (± 3 days) Assessments[1]	Weekly Phone Assessment (Weeks 2, 3 and 6 , (± 3 days)	Unscheduled Possible Recurrence Visit	Unblinded Enema Administration[1] (≤10 calendar days from Tx Failure)	Phone Assessment at 3 and 6 months (± 14 days)
Informed consent obtained	X						
Demographics, medical history	X						
Prescribe/continue antibiotics for CDI symptom control	X						
Modified physical exam conducted		X					
Stool sent to Rebiotix by subjects for testing and archiving (optional)	X		X		X		X
Central Lab CBC w/differential testing[3]	X[3]	X					
Central Lab CMP & CRP testing	X						
C. difficile testing[4]	X[4]				X[4]		
Central Lab stool and blood testing		X					
Urine pregnancy testing performed at site (if applicable)	X	X				X	
Cdiff32 Questionnaire	X		X		X		X
Inclusion/exclusion criteria confirmed	X	X					
Randomization assignment	X[5]						
24-72hr washout period confirmed		X				X	
Enema administered		X				X	
Product complaint (if applicable)		X				X	
Recurrence or New CDI symptoms assessed		X	X	X	X	X	X
Vital signs assessed	X	X			X	X	
Subject Diary discussed/reviewed	X	X	X[6]			X	
Employment status assessed	X		X[7]				X
Concomitant medications	X	X	X	X	X	X	X
Medical History Assessment	X		X[7]		X		X
Adverse events assessed		X	X	X	X	X	X
Solicited events assessed[2]		X	X[2]			X[2]	
Protocol deviations (if applicable)	X	X	X	X	X	X	X

[1] Documented treatment failures may receive an unblinded RBX2660 enema. If an unblinded enema is administered, the follow-up visit requirement re-start based on the date of last enema administration.
[2] Solicited events are collected in the Subject Diary from the day of administration of any study enema (blinded or unblinded) until the day prior to the 1-week visit. The Subject Diary is collected and reviewed at the 1-week visit. Solicited events that increase in severity from screening should be assessed for a possible adverse event.
[3] Exclusion criteria for absolute neutrophil count should be assessed based on the CBC collected at the screening visit.
[4] Perform within 30 days prior to or at enrollment and if CDI recurrence is suspected.
[5] Randomization should occur after screening criteria has been assessed and eligibility confirmed (4-14 calendar days from randomization to blinded study treatment)
[6] Subject Diary is reviewed only at the 1 week follow-up visit
[7] Collected at the 8 week follow-up visit only.

6.2 Time Point Algorithms

6.2.1 Relative Day

The date of study enema administration will be considered relative day 1, and the day before study enema administration will be relative day -1. Relative days will be calculated as follows only when the full assessment date is known (i.e., partial dates will have missing relative days, unless imputation methods have been specified):

For days before first enema:
Date of Assessment – Date of first enema administration

For days on or after first enema:
Date of Assessment – Date of first enema administration +1

6.2.2 Windows

For the purpose of statistical analysis, the analysis visit windows will be calculated in terms of study days since the day of the study enema administration, as illustrated in the following tables:

Analysis Windows for Clinical Visits

Visit	Scheduled Study Day	Visit Window in Protocol (Days)	Visit Window for Analysis (Days)
Screening Visit	-30 through -1	-30 through -1	-30 through -10
Baseline/1st Enema	1	1	1
Week 1 Office	8	5 - 11	5 - 11
Week 4 Office	29	26 - 32	26 - 32
Week 8 Office	57	54 - 60	54 - 60

Analysis Windows for Phone Assessments

Visit	Scheduled Study Day	Visit Window in Protocol (Days)	Visit Window for Analysis (Days)
Week 2 - Phone	15	12 - 18	12 - 18
Week 3 - Phone	22	19 - 25	19 - 25
Week 6 - Phone	43	40 - 46	40 - 46
3 month - Phone	90	76 – 104	76 – 104
6 month - Phone	180	166 – 194	166 – 194

If a subject has more than one assessment occurring in the same visit window, the data from the visit closest to the scheduled study day will be used for analysis. If two visits have the same distance from the scheduled study day, i.e., one visit occurs -1 day from study day, and the other visit occurs +1 day from study day, the data from the visit after the scheduled study day will be used for analysis.

Analysis Windows for Adverse Event Reporting by Onset Interval

Onset Interval Name	Onset Interval Analysis Window	Window for Analysis (Days)
Baseline	Screening - < Baseline	≤ -1
1-week	Baseline - 1-week	1 through 8
4-week	> 1-week - 4-week	9 through 29
8-week	> 4-week - 8-week	30 through 57
4-month	> 8-week - 4-month	58 through 121
6-month	> 4-month - 6-month	122 through 180+

6.3 Analysis Populations

The modified Intent-to-Treat will be analysis of import for the primary endpoint.

6.3.1 Intent-to-Treat Population

Analysis of the primary efficacy endpoint will be done using the Intent-to-Treat (ITT) population as a sensitivity analysis. The ITT population is defined as all randomized subjects. Subjects will be analyzed according to the randomized treatment rather than the actual treatment received regardless of treatment misallocations. Re-randomized subjects will be summarized using their second treatment assignment (see Section 4.3 for details). Randomized subjects who exited prior

to receiving blinded treatment will not be included in the analysis. Randomized subjects in whom blinded treatment was attempted but delivery of the enema was not successful will be counted as treatment failures. Randomized and treated subjects who exited prior to their 8-week efficacy assessment will be counted as treatment failures even if protocol required treatment failure documentation is not on file. Additionally, the secondary efficacy endpoints may be evaluated using the ITT population.

6.3.2 Modified Intent-to-Treat Population

For the Modified Intent-to-Treat (mITT) population, subjects will be analyzed according to the randomized treatment assignment despite any treatment misallocations.

Additionally, the mITT population will be defined as all randomized subjects who successfully received blinded treatment but excluding;

- subjects who withdrew prior to treatment;
- subjects in whom treatment was attempted but not completed and;
- subjects who discontinue from the study prior to evaluation of treatment failure/success for the primary endpoint if the reason for exit is not related to CDI symptoms. Reason for exit will be captured on the exit form to allow for identification of such subjects.

Examples of reasons unrelated to CDI symptoms may include;

- Withdrawal of consent
- Death unrelated to CDI

Subjects who were exited prior to the 8-week efficacy assessment due to CDI related symptoms will be counted as treatment failures even if documentation as required per the protocol is not on file for the treatment failure.

The mITT population will be used for the primary efficacy reporting purposes. The mITT population will also be used for secondary and exploratory endpoints evaluation. The analysis of the ITT and PP populations may also be provided in addition to the mITT for purposes of regulatory evaluation.

6.3.3 Per-Protocol Population

The per-protocol (PP) population will consist of all subjects who successfully received blinded treatment analyzed according to the treatment they received, excluding;

- Subjects who have documented deviations to inclusion or exclusion criteria.
- Subjects who exited prior to the 8-week efficacy evaluation if the reason for exit was not related to CDI symptoms in the same manner as the mITT population.

The primary, secondary and exploratory endpoints will be evaluated using the PP population.

6.3.4 Safety Population

The safety population (SP) will be defined as the population of randomized subjects who had any blinded treatment attempted or completed. Subjects will be analyzed according to the treatment they actually received should misallocations occur. The safety population will be used in analysis of all safety endpoints.

6.4 Efficacy Variables

6.4.1 Primary Efficacy Endpoint

The primary efficacy endpoint is the recurrence of CDI within 8 weeks of blinded treatment. The primary analysis of the study will be a Bayesian hierarchical model, which formally incorporates data from a previous randomized Phase 2B study (Protocol 2014-01) of RBX2660. This analysis will demonstrate the efficacy of RBX2660 by showing a sufficiently high posterior probability that the rate of success in the treatment group is superior to the rate of success in the control group.

6.4.2 Secondary Efficacy Endpoint

The secondary efficacy endpoint is the recurrence of CDI through 6 months after blinded treatment. This will be analyzed as the difference in time to CDI event between RBX2660 and Placebo following blinded treatment through 6 months of follow-up. This analysis will only be

analyzed if statistical significance ($p<0.05$) is demonstrated for the primary efficacy endpoint. These endpoints will be tested at a two-sided 0.05 significance level.

The null hypothesis of interest is that there is no difference between RBX2660 and Placebo treated subjects in the probability of CDI at any time point during the 6 months of follow-up. The alternative is that there is a difference in the probability of CDI at any time point during 6 months of follow-up.

The power to detect the difference between RBX2660 and Placebo treatment at 6 months was estimated by 10,000 simulations of time to event data, with 104 subjects per treatment group. The data was simulated assuming the exponential distribution with the rate parameter lambda estimated using the 8-week success rates of 69% and 47% for RBX2660 and Placebo, respectively. Using the log-rank test, the power to detect the difference between the two groups at 6 months was 99.5%.

6.4.3 Other Efficacy Endpoints

Other efficacy endpoints include the following:

- Baseline characteristics
- Subject fecal microbial composition at Screening, and 4-weeks, 8-weeks, 3 months and 6 months after blinded study treatment.
- Charlson Comorbidity Index at Screening, 8-week, 3-, and 6-month phone assessments.
- Cdiff32 questionnaire at Screening, Week 1, 4, 8, month 3 and 6
- ATLAS score for CDI severity of qualifying CDI event.
- Recurrence of CDI within 8 weeks of unblinded (open-label) RBX2660 treatment in Placebo subjects who were documented blinded study treatment failures.
- Recurrence of CDI within 8 weeks of unblinded (open-label) RBX2660 treatment in subjects who received a blinded dose of RBX2660.
- Occurrence of CDI through 6 months in all subjects receiving a single dose of RBX2660.
- Concentration of vancomycin resistant enterococcus in stool samples for subjects who were carriers at baseline.
- Presence of *C. difficile* in stool samples at Screening, and 4-weeks, 8-weeks, 3 months and 6 months after blinded study treatment.

6.5 Safety Variables

All the safety analyses will be based on the safety population; see Section 6.3.4 in this document.

6.5.1 Safety Endpoints

- Number of adverse events per subject
- Timing of attributable adverse event post-treatment exposure (treatment emergent adverse event TEAE)
- Duration of TEAE
- Relatedness of TEAE
- Severity of TEAE
- Causality of TEAE to IP, enema, *C. difficile* or prior condition Number of each of the following through 8 weeks post blinded treatment: death, septic shock, toxic megacolon, colonic perforation, emergency colectomy, and ICU admission.
- Onset of new chronic conditions relative to blinded treatment administration.

6.5.2 Extent of Exposure to Study Treatment

The number and proportion (RBX2660 vs Placebo) of enemas administered for each subject will be summarized using descriptive statistics by treatment arm at each study enema administration (blinded and unblinded). Time between study enema administrations will be summarized using descriptive statistics for both blind and unblinded enemas. Additional analysis will be done for those subjects who received both Placebo and RBX2660 enemas.

6.5.3 Adverse Events

An adverse event (AE) is any untoward medical occurrence associated with the use of IP, whether or not the event is considered product-related. Preexisting conditions that worsen in frequency, intensity or character of the condition during study participation will be recorded as an AE. AEs will be recorded for each subject from the day of enrollment through the 6-month

telephone assessment. Serious adverse events (SAEs) will be recorded for each subject from the day of enrollment through the duration of the study.

6.5.3.1 Adverse Event Dictionary

AEs and SAEs will be classified by the current Medical Dictionary for Regulatory Activities (MedDRA) version. The verbatim term recorded by the investigator will be mapped in MedDRA and system organ class (SOC), high level group term (HLGT), high level term (HLT), preferred term (PT), and lower level term (LLT) will be attached to the clinical database. Tables and listings will present data at the SOC and PT level.

6.5.3.2 Adverse Event Severity

Severity of AEs will be graded as mild, moderate, severe, or potentially life-threatening. The severity grade of events for which the investigator did not record severity will be categorized as "Unknown" for tabular summaries and data listings, and will be considered the least severe for the purposes of sorting for data presentation. For summaries by severity grade, if a subject has multiple events occurring in the same SOC or PT, then the most severe event will be selected.

6.5.3.3 Relationship of Adverse Events to Study Treatment

Relationship of an AE to the Investigational Product (IP), the enema procedure, *C. difficile* disease, and a preexisting condition will be classified by the investigator as definite, probable, possible, or unrelated. Related events will include AEs that are classified as definitely, probably, or possibly related. IP includes both the Placebo and RBX2660 enemas.

6.5.3.4 Treatment-Emergent Adverse Events

Adverse events will be considered treatment-emergent adverse events (TEAE) according to the following algorithm:

- If the complete onset date of an AE is known, then:
    - If AE onset date is prior to initial treatment date, then the AE will not be considered treatment-emergent.

- If AE onset date occurs on or after initial treatment date, then the AE will be considered treatment-emergent.

- If AE onset date is partially known, then:
  - If day is unknown, and month/year occurs on or after first enema month/year, then AE will be considered treatment-emergent. Otherwise, AE will not be considered treatment-emergent.
  - If month/day is unknown, and year occurs on or after first enema year, then AE will be considered treatment-emergent. Otherwise, AE will not be considered treatment-emergent.

6.5.4 Other Observations Related to Safety

6.5.4.1 Solicited Adverse Events

The following list of anticipated adverse events are solicited from subjects via the Subject Diary from the date of enrollment through the 7th day after receiving the randomized study treatment. For those subjects who receive an unblinded RBX2660 enema, these events are again collected in a new post-treatment Subject Diary from the day of the unblinded RBX2660 enema to the 7th day after receiving the unblinded RBX2660 enema:

- gas (flatulence)
- abdominal distension or bloating
- rectal irritation or pain
- chills/severe shivering
- abdominal pain or cramping
- increased diarrhea
- constipation
- rectal bleeding
- nausea
- vomiting
- fever ≥ 37.8° C (≥100.0°F)

Frequency and severity of the solicited AEs will also be captured in the Subject Diary.

6.5.4.2 Major Complications of CDI Events

Frequencies of major complications of CDI including death, septic shock, toxic megacolon, colonic perforation, emergency colectomy, or ICU admission will be collected through the 6-month telephone assessment.

6.5.4.3 Hospitalizations for Recurrent CDI

Details regarding hospitalizations due to CDI, including admission and discharge date/time, reason for admission, whether the subject was admitted to an ICU, and ICU admission and discharge date/time (if applicable), will be collected following both the blinded and unblinded study enemas.

7 STATISTICAL ANALYSIS

7.1 Disposition of Subjects

A summary table will present subject disposition for all subjects enrolled. The table will show the number of subjects enrolled (consented), the number of enrolled subjects not randomized, followed by the number of subjects randomized for each treatment group and lastly the number of subjects randomized and treated. The proportion of randomized subjects in each treatment group and overall who complete the study and who discontinue the study will be presented. Separately, a table or listing will be created to document the subjects that are included in each analysis population; ITT, mITT, PP and SP.

Discontinuation will be categorized by reason as a percentage of the number of subjects in each analysis population and overall. Additionally, subjects who were successfully treated but withdraw prior to the 8-week primary endpoint evaluation will be identified along with the reason for withdrawal to document which analysis population(s) to which they contribute.

7.2 Protocol Deviations

Protocol deviations will be documented in the clinical database. The information available will be summarized by treatment arm and deviation type and listed by individual site.

7.3 Baseline Assessments

Baseline assessments are collected within 30 days before or on the day of study treatment administration. Summaries for age by group (< 65 and ≥ 65), sex, ethnicity, and race will be provided by analysis group and treatment arm. Baseline height, weight and employment status will also be summarized by analysis group and treatment arm. CDI history will be summarized by total number of episodes experienced prior to blinded treatment. Additionally, the following will be summarized for each episode of CDI: duration (days), treatment administered, hospitalization with duration (if applicable) and available *C. difficile* test results. Additionally, the severity of the qualifying CDI episode will be documented for each subject by ATLAS score and the baseline comorbidities will be scored for each subject using the Charlson Comorbidities Index.

The following baseline assessments will be provided in summary table and listings by treatment arm:

- Medical history
- CDI history
- Laboratory testing (hematology)
- Vital signs (systolic/diastolic blood pressure [mmHg], pulse rate [beats/min], respiration rate [breaths/min], temperature [°C])

7.4 Demographic and Other Baseline Characteristics

All baseline summaries will include the ITT, mITT and PP population and will be provided by treatment arm. Descriptive statistics (number of subjects [N], mean, standard deviation [SD], median, interquartile range [IQR], minimum, and maximum for continuous variables and number

of subjects [N] and the percentage for categorical variables) will be provided for all baseline measures.

7.5 Prior and Concomitant Therapy

The current World Health Organization (WHO) Drug Dictionary will be used to classify medications by preferred term and WHO Anatomical Therapeutic Chemical (ATC) at the highest level of ATC. Listings of all concomitant medications will be made by WHO ATC classification of ingredients and by preferred term tabulated by each treatment arm. All listings of usage of medications will be based on the Safety population.

7.6 Analysis of Efficacy

7.6.1 Analyses Supporting the Primary Efficacy Endpoint

7.6.1.1 Treatment Success

A hierarchical, closed-testing procedure will be utilized for the primary endpoint and secondary endpoints. Analyses of the single primary endpoint, recurrence of CDI within 8 weeks of blinded treatment, will be analyzed as the proportion of treatment success of Group A Placebo (P_{PLA}) and Group B RBX2660 (P_{RBX}). This analysis will be performed on the ITT, mITT and PP populations, with the mITT being the analysis of import for reporting purposes.

The primary efficacy analysis will be a Bayesian hierarchical model which formally incorporates data from the previous Phase 2B study (Protocol 2014-01) of RBX2660. Study success will be declared if the posterior probability of superiority exceeds 0.99943 at the first two interim analysis. If the trial continues to the maximum enrollment, the success criteria will be chosen so that the overall alpha is a nominal 0.00125. Success at this level is considered appropriate for approval based on results from a single study. If success at this level is not met at the final analysis, success may also be declared if the posterior probability exceeds 0.97706. Success at this secondary threshold still constitutes a successful study, although it may not be sufficiently strong evidence to request regulatory approval based on a single study without further FDA discussion. Details regarding the analysis can be found in Section 10 (Bayesian Analysis Plan).

The secondary endpoints will be analyzed using a hierarchical testing approach if sufficient evidence is found to claim study success.

The primary efficacy analysis will be conducted when all subjects in each group reach their 8-week assessment following blinded treatment or early termed from the study. Full study results in the form of a final analysis and report will be completed once the last subject has completed the 6-month telephone assessment, at which time the study will be considered complete. This includes subjects receiving unblinded RBX2660 due to a documented treatment failure.

7.6.1.2 Sensitivity Analysis for the Primary Efficacy Endpoint

The primary efficacy analysis will be repeated using the ITT and PP population to assess the sensitivity of the primary endpoint regardless of the outcome of the mITT primary efficacy analysis. The ITT analysis will serve as the conservative analysis as it considers all subjects who exit prior to the 8-week efficacy assessment as treatment failures regardless of treatment failure documentation.

Three groups of sensitivity analyses will be performed to assess the sensitivity of the results to the borrowing of historical data from the previous Phase 2B study (Protocol 2014-01) of RBX2660.

1) A Pearson's chi-square test will be used to test the null hypothesis that the response rate in the treatment group is equal to that of the control group. In addition, two-sided 95% confidence intervals for the difference in response rate between arms will be calculated using a normal distribution approximation. This sensitivity analysis will be conducted using the mITT, ITT, and PP populations using data from the 2017-01 study only.
2) Sensitivity of the primary analysis (Bayesian analysis) to the analysis population (ITT, PP) using Bayesian analysis.
3) Sensitivity of the primary analysis (Bayesian analysis) to the following different sensitivity analysis populations from the 2014-01 study using Bayesian analysis. The **2014-01 Sensitivity Analysis Population #1** represents the ITT population from the 2014-01 study, and the **2014-01 Sensitivity Analysis Population #2** represents the

population used for the primary analysis inclusive of one subject in Group C who was exposed to product but did not receive the full enema.

2014-01 Sensitivity Analysis Population #1				**2014-01 Sensitivity Analysis Population #2**			
2014-01 Treatment Group	# Subjects	# Responders	Response Rate	2014-01 Treatment Group	# Subjects	# Responders	Response Rate
A	45	25	0.556	**A**	41	25	0.610
B	44	19	0.432	**B**	44	19	0.432
C	44	25	0.568	**C**	43	25	0.581

Efficacy analysis will be repeated without the borrowing of historical data from the Phase 2B study (Protocol 2014-01) using the mITT, ITT, and PP populations.

An analysis of time to CDI recurrence after completion of the assigned blinded treatment will be analyzed and presented by the Kaplan-Meier procedure and median duration presented by days (and weeks) using the ITT, mITT, and PP populations. The Kaplan-Meier estimate of the median time (days) and the associated 95% confidence interval will be presented for each treatment group. A graphic presentation of the Kaplan-Meier estimates will supplement the tabular presentation. Between group comparisons will be performed using the stratified log-rank test based on the stratification factor at randomization. The corresponding hazard ratio of the treatment effect along with the 95% confidence interval will be calculated using a Cox proportional hazard model with treatment and the stratification factor at randomization as explanatory variables.

Time to CDI recurrence is defined as the number of days from enema administration to first assessment indicating recurrence for those subjects who were deemed Treatment Failures with CDI recurrence. All subjects who are Indeterminate or discontinued for any reason prior to assessment of efficacy will be censored at the last assessment date at/prior to Week 8. For the analysis of Time to CDI recurrence within 8 weeks, all subjects considered as a treatment success will be censored at the date of their 8-week assessment. For the analysis of Time to CDI

recurrence through 6 months, all subjects considered to have achieved sustained clinical response will be censored at the date of their 6-month assessment.

Finally, the sensitivity of the primary efficacy endpoint will be assessed in a multivariate logistic regression adjusting for possible interactions of the following covariates with treatment group: age (< 65 years, ≥ 65 years), sex (Female, Male), race group (White, Non-white), ethnicity (Hispanic-Latino, not Hispanic-Latino), site geography (outside the US, Eastern US*, Southern US*, Northern US*, Western US*), randomization strata (i.e., treatment received at the qualifying event) and number of previous episodes of CDI recurrence at baseline. Other baseline variables will be compared for balance using the appropriate statistical test. This analysis will be performed on all analysis populations where at least 20 subjects in each sub-group are available for analysis to fully understand the observed treatment effect.

*States included in the identified US geographical regions are as follows:

- Eastern US: Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, Vermont, Virginia, West Virginia
- Southern US: Alabama, Arkansas, Georgia, Florida, Kansas, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas
- Northern US: Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, Wisconsin
- Western US: Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming

7.6.2 Analysis Secondary Efficacy Endpoint

A hierarchical, closed-testing procedure will be utilized for the secondary endpoint: If the null hypothesis for the primary efficacy endpoint is rejected, the secondary endpoint will be analyzed. The secondary endpoint will be tested at a two-sided 0.05 significance level.

7.6.2.1 Sustained clinical response of RBX2660 vs. Placebo.

The null hypothesis of interest is that the rates of sustained clinical response (i.e., the rate of CDI occurrence with treatment success of the presenting CDI recurrence and no new CDI episodes for greater than 8 weeks after completing a blinded study treatment) are the same between the RBX2660 group and the Placebo group during the 6 months of follow-up. The alternative is that there is a difference in sustained clinical response rates. A Pearson's chi-square test will be used to test the null hypothesis that the response rate in the treatment group is equal to that of the control group. In addition, two-sided 95% confidence intervals for the difference in response rate between arms will be calculated using a normal distribution approximation. The null hypothesis will be rejected at the two-sided 0.05 significance level. Subjects who were deemed a Treatment Success for the efficacy analysis and exited prior to their 6 month follow-up will be conservatively counted as a sustained treatment failure for the purpose of this analysis.

An analysis of time to CDI occurrence after completion of the assigned blinded treatment will be analyzed and presented by the Kaplan-Meier procedure and median duration presented by days (and weeks) using the ITT, mITT, and PP populations. The estimates will be presented for each of the time points at which CDI recurrence assessments are scheduled to occur. A graphic presentation of the Kaplan-Meier estimates will supplement the tabular presentation. Between group comparisons will be performed using the log-rank test.

Time to CDI occurrence is defined as the number of days from enema administration to first assessment indicating recurrence, for those subjects who were deemed Treatment Failures or Indeterminate. Randomized subjects who do not complete the assigned blinded treatment will be censored at Day 0. All subjects who are discontinued for any reason prior to the 6-month timepoint will be censored at the last assessment date at/prior to Month 6. All subjects considered as a sustained treatment success will be censored at the date of their 6-month assessment.

7.6.3 Other Endpoint Analysis

An initial assessment will be completed during the primary efficacy analysis as follow-up data allows with a complete analysis to be provided with the full study results for applicable other analyses. Additionally, the other analyses will be conducted on other efficacy populations as needed.

7.6.3.1 Comparison of baseline characteristics in subjects successful with one dose of RBX2660 vs. those who were unsuccessful.

A multivariate analysis or regression analysis will be used to identify predictors (if any) for subject success with one RBX2660 dose vs. those who were unsuccessful.

7.6.3.2 Cdiff32 scores at Screening, 1-week, 4-week, and 8-week assessment visits, and 3 and 6 month phone assessments.

Changes from Screening in the Cdiff32 scores will be summarized using descriptive statistics by treatment arm. Last observation carried forward (LOCF) will be used to impute missing post-baseline data. No inferential statistics will be performed.

7.6.3.3 Subject fecal microbial composition at Screening, and 4 and 8- weeks and 3 and 6 months after blinded study treatment.

Descriptive statistics will be used to present the changes from Screening fecal microbial composition to those at 4 weeks, 8 weeks, 3 months and 6 months after blinded study treatment for those subjects treated with RBX2660 and those receiving Placebo.

7.6.3.4 Charlson Comorbidity Index at Screening, 8 week visit, and 3-, 6- month phone assessments.

A comorbidity index score will be documented for all subjects at Screening and compared to subsequent scores at 8 weeks, 3 months and 6 months. An initial assessment will be completed during the primary efficacy analysis as follow-up data allows with a complete analysis to be provided with the full study results. Additional analysis may be done of Placebo subjects who

were documented treatment failures after blinded treatment then treatment successes following unblinded RBX2660 treatment.

7.6.3.5 Atlas score for CDI severity of qualifying CDI.

Quantification of CDI severity scores for subjects at the time of the CDI recurrence that qualified them for enrollment into the study based on the ATLAS Scoring System. Data from the qualifying CDI event will be used (see table below), including the protocol required eligibility criteria of treatment with systemic antibiotics prior to randomization. However, if data points are not available, baseline data may be substituted.

Parameter	0 Points	1 Point	2 Points
Age (years)	< 60	60 – 79	≥ 80
Temperature (°C)	≤ 37.5	37.6 – 38.5	≥ 38.6
Leukocyte count (total)/mm^3	< 16,000	16,000 – 25,000	> 25,000
Albumin (g/L)	> 35	26 – 35	≤ 25
Systemic concomitant antibiotics during CDI RX	No	-	Yes

7.6.3.6 Recurrence of CDI within 8 weeks in Placebo subjects who were documented study treatment failures and then received RBX2660 through open-label.

Treatment success for documented Placebo failures after receiving RBX2660 will be completed using a Chi-square test comparing the treatment success of blinded placebo subjects as compared to the success of placebo subjects who received RBX2660. This analysis will be completed with the full study results rather than the primary efficacy analysis as it requires data on the outcome following any unblinded RBX2660 enemas received.

7.6.3.7 Recurrence of CDI within 8 weeks of unblinded (open-label) RBX2660 treatment in subjects who received a blinded dose of RBX2660.

Determine the treatment success rate of those subjects randomized to RBX2660 who were deemed an initial treatment failure then went on to receive an unblinded RBX2660 enema (two RBX2660 enemas). This success rate will be calculated at 8 weeks following the second

RBX2660 enema and compared to the blinded success rate of the RBX2660 arm using a Chi-square test.

7.6.3.8 Occurrence of CDI through 6 months in all subjects receiving a single dose of RBX2660.

Determine the combined treatment efficacy for RBX2660 randomized subjects plus Placebo randomized subjects who were documented failures after blinded treatment then received a unblinded RBX2660 enema. This analysis will be completed with the full study results rather than the primary efficacy analysis as it requires data on the outcome following any unblinded RBX2660 enemas received. Both treatment success and treatment durability will be analyzed;

- Treatment success rate for the prevention of CDI recurrence
- Treatment success rate for the prevention of new CDI episodes (those with onset occurring beyond 8 weeks after study treatment)

7.6.3.9 Concentration of vancomycin resistant enterococcus in stool samples for subjects who were carriers at baseline.

Lab generated values for the presence of vancomycin resistant enterococcus based on submitted stool samples at baseline and at 1-, 4-, 8-weeks, and 3 and 6 months will be assessed for all subjects to identify any trends in the clearance of vancomycin resistant enterococcus following enema treatment.

7.6.3.10 Presence of C. difficile in stool samples at Screening, and 4-weeks, 8-weeks, 3-months and 6-months after study treatment.

Lab generated values for the presence of *C. difficile* based on submitted stool samples at baseline and at 1-, 4-, 8-weeks, and 3 and 6 months will be assessed for all subjects to identify any trends in the clearance of *C. difficile* following enema treatment.

7.7 Safety Analysis

The Safety population will be used to summarize all adverse event data, unless otherwise specified. Safety data will be summarized separately, using the same approach, for both the blinded and unblinded enemas. Adverse events will be summarized for each randomization group through the 8-week primary endpoint follow-up. Adverse events occurring after the 8-week follow-up visit will be grouped into the following categories:

1. Group A, placebo randomized subjects who did not receive an unblinded RBX2660
2. Group B, RBX2660 randomized subjects who did not receive an unblinded RBX2660
3. Group C, placebo randomized subjects who went on to receive an unblinded RBX2660
4. Group D, RBX2660 randomized subjects who went on to receive an unblinded RBX2660

7.7.1 Adverse Events

Adverse events will be summarized using counts of the number of events and frequencies (counts and percentages) of the number of subjects by treatment group and overall. The incidence of each preferred term (PT) within the primary system organ class (SOC) as well as overall primary SOC incidence will be presented. Summaries may also be presented by PT, irrespective of SOC. A subject with multiple events coded to the same PT within a primary SOC will be counted only once for the PT within the primary SOC. Likewise, a subject with multiple events coded to the same SOC will be counted only once within the SOC.

An overall summary of AEs will present the incidence of (1) all TEAEs, (2) TEAEs related to IP, (3) TEAEs related to enema procedure, (4) TEAEs related to *C. difficile* disease, (5) TEAEs related to a preexisting condition, (6) serious TEAEs, (7) serious TEAEs related to IP, (8) serious TEAEs related to the enema procedure, (9) serious TEAEs related to *C. difficile* disease, (10) serious TEAEs related to a preexisting condition, and (11) TEAEs leading to death.

Frequency of occurrence and number and percent of subjects experiencing an event through 6 months follow-up will be presented by treatment for both the blinded and unblinded enemas as follows:

- All TEAEs
- TEAEs by maximum severity
- TEAEs by relatedness to IP
- TEAEs by relatedness to *C. difficile* disease
- TEAEs by relatedness to enema procedure
- TEAEs by relatedness to preexisting condition
- TEAEs resulting in subject withdrawal
- TEAEs resulting in death
- Serious TEAEs
- Serious TEAEs by maximum severity
- Serious TEAEs by relatedness to IP
- Serious TEAEs by relatedness to *C. difficile* disease
- Serious TEAEs by relatedness to enema procedure
- Serious TEAEs by relatedness to preexisting condition
- Serious TEAEs resulting in subject withdrawal
- Serious TEAEs resulting in death

For the summary of AEs by maximum severity, if a subject has multiple events occurring in the same SOC or same PT, then the event with the highest severity will be counted. For AEs reported by causality, if a subject has multiple events occurring in the same SOC or same PT, the event with the highest association will be summarized.

A summary of TEAE and serious TEAE characteristics (severity, relatedness to IP, relatedness to *C. difficile* disease, relatedness to enema procedure, and relatedness to preexisting condition by outcome) will also be presented by treatment and overall for both the blinded and unblinded enemas.

Incidence of AEs and serious AEs will also be presented by onset interval after the blinded and unblinded enemas for the following intervals:

- Baseline (onset date prior to date of blinded enema);

- IP administration (onset date on the date of the blinded or unblinded enema as applicable);
- and the following intervals relative to the last enema date (blinded or unblinded as applicable):
    - 1 Week (day 8), 4 Weeks (day 29), 8 Weeks (day 57), 3 Months (Day 90) and 6 Months (Day 180).

At the time of the primary efficacy analysis, all safety endpoints related to events occurring within 8 weeks following blinded treatment will be assessed. Additionally, other safety endpoints will be analyzed as the data allows during the primary efficacy analysis with a complete analysis to be provided with the full study results in the final analysis.

For data presentation, SOC will be ordered alphabetically, with PT sorted by decreasing total frequency. No inferential statistics will be provided for the AE data.

Subject listings will be presented for all AEs including SAEs.

7.7.2 Other Analyses Related to Safety

7.7.2.1 Subject Diary including Solicited Adverse Events and Compliance

All solicited adverse events recorded in the Subject Diary will be summarized utilizing frequencies (number of events), proportions (number of subjects with the event) and 95% CIs for categorical variables and using the descriptive statistics as listed in Section 8.1 for continuous variables.

For each solicited event, severity scores will be summarized as both categorical and continuous values. For continuous summaries, severity scores will be identified as either pre-treatment or post-treatment (all diary entries captured after the blinded and unblinded enemas), and the average for each period will be calculated across individual subjects. Baseline for continuous summaries will be considered the average during the pre-treatment period. For categorical summaries, if a subject reports more than one severity score on the same date, the maximum

severity will be used. Baseline for categorical summaries will be considered the last non-missing assessment prior to the blinded enema.

For each categorical summary of event frequency, a two-sided 95% CI will be calculated using the normal approximation of the binomial.

The following summaries will be produced separately for the blinded enemas and unblinded enemas:

- Frequency of severity scores on day of enema administration by treatment and overall
- Frequency of severity scores on day 1 through 1-week post-treatment by treatment and overall
- Frequency of maximum post-treatment severity score by treatment and overall

The following summaries will be produced for the double-blind phase only:

- Summary of average severity score and change from baseline by treatment and overall
- Shift from baseline to 1-week post-treatment severity score by treatment and overall
- Shift from baseline to maximum post-treatment severity score by treatment and overall

The following summary will be produced during the final analysis as related to the overall scores:

- Summary of average severity scores by treatment and overall

Compliance in returning the Subject Diary will be calculated by dividing the number of diaries turned in, whether complete or not, by the number of diaries expected to be turned in. Furthermore, compliance will be summarized by calculating the proportion of subjects who turned in "complete" diaries, where complete is defined as answering all 11 daily questions for at least 80% of the expected number of diaries. Compliance will be calculated following both the blinded and unblinded enemas separately and summarized by treatment group.

7.7.2.2 Major Complications of CDI events

Major complications of CDI events include death, septic shock, toxic megacolon, colonic perforation, emergency colectomy, or ICU admission collected through the 6-month telephone assessment. Incidence of major complications of CDI events will be summarized by treatment arm after both the blinded and unblinded enemas.

7.7.2.3 Hospitalizations for Recurrent CDI

The number of subjects requiring hospitalization due to recurrent CDI through 8 weeks following the completion of the randomized study treatment will be summarized, along with the length of hospital stay, the number of subjects requiring ICU admission, and the length of ICU stay during hospitalization.

A listing of all hospitalizations due to CDI throughout the study will be generated.

7.7.2.4 Rate of onset of new chronic conditions comparing those treated with RBX2660 vs. Placebo.

Documentation of physician diagnosed chronic conditions for all subjects will be completed at baseline and compared to new physician diagnosed chronic conditions reported at 8 weeks, 3 months, and 6 months. Chronic condition is defined as a condition lasting >3 months. An initial assessment will be completed during the primary efficacy analysis as follow-up data allows with a complete analysis to be provided with the full study results. Additional analysis may be done of Placebo subjects who were documented treatment failures after blinded treatment then treatment successes following unblinded RBX2660 treatment.

8 STATISTICAL METHODS

8.1 General Methodology

All statistical tests will be two-sided with a significance level of α=0.05, unless specified otherwise, and will be performed using SAS® Version 9.3 or higher. Confidence intervals (CI) will be constructed at the level of 95% unless specified otherwise. For continuous variables,

descriptive statistics (number of subjects [N], mean, standard deviation [SD], median, inter-quartile range [IQR], minimum, and maximum) will be generated. For discrete/categorical variables, the number and percentage of non-missing subjects will be generated. Descriptive statistics will be provided by treatment arm for all subjects. Standard operating procedures (SOPs) will be followed in the creation, validation and quality control of all data displays and analyses.

Subject listings of all data from the eCRFs as well as any derived variables will be presented.

8.2 Adjustments for Covariates

Adjustments for covariates will be made in sensitivity analyses of the primary endpoint if applicable due to imbalance of a particular covariate.

8.3 Handling of Dropouts or Missing Data

No imputations will be used in analyses/summaries. However, a last observation carried forward analysis will be used for evaluation of the Cdiff 32 scores. Missing safety data will not be imputed beyond the adjustments for adverse event data mentioned in Section 6.5.2.

8.4 Medical Monitor Adjudication

The Medical Monitor reviews and adjudicates all serious adverse events and other adverse events of interest to provide an objective, qualified judgment of the events. Adverse event review and adjudication includes the event description, onset date in relation to the treatment date, and the investigator's determination of causal relatedness to IP, the enema procedure, and pre-existing condition, as well as rating the event's severity and seriousness. The Medical Monitor will remain blinded throughout the study. However, the DSMB will be unblinded as needed to adjudicate specific adverse events as per the study stopping rules (see Section 8.5.1, below).

8.5 Efficacy Analyses and Data Monitoring

The EAC will adjudicate individual efficacy data at the time of the interim and final analyses. Interim analyses evaluating the primary efficacy endpoint are planned once 160 subjects (minimum) and 220 subjects in the mITT population have completed (i.e., completed the 8 week follow-up assessment and/or completion of the Treatment Outcome form). At each interim analysis, the study may declare early success, in which case enrollment is stopped and the sponsor will file for immediate approval. The study may also declare futility, resulting in the cessation of enrollment and follow-up.

A primary efficacy and safety analysis will be performed when 100% of subjects reach their 8-week assessment following blinded treatment or are documented treatment failures by the site investigator. Safety analysis will be completed for safety endpoints defined at or prior to 8 weeks following blinded treatment. Full safety results as well as unblinded RBX2660 efficacy data will be reported after all available subjects complete their last 6-month telephone assessment following their last study enema (blinded or unblinded).

To decrease the chance of introducing bias, the primary efficacy analyses will use efficacy data as reviewed and confirmed by the EAC. The study sites, subjects, and the medical monitor will remain blinded to the specific results and to individual randomized treatment assignments until the study is completed.

8.5.1 Study Stopping Rules

Enrollment will be paused if any of the following events are identified and the DSMB determines that there is probable cause that the IP or enema procedure contributed to the event:

1. There is probable cause that IP or enema procedure (e.g., due to transfer from an RBX2660 donor) contributed to a pathogenic intestinal infection in the stool of any subject, or

2. Any series of events of major significance such as death or other serious outcome for which a causal connection with the IP is plausible represents an excess of the important adverse event(s) in one of the study arms.

8.6 Multi-center Studies and Pooling of Centers

The study will be conducted at up to 80 clinical study sites in the US and Canada.

This is a multicenter study and before data from all sites are pooled, the effect of site will be assessed for the primary efficacy endpoint. A sufficient number of subjects (≥ 7 meeting the PP definition) will be required within each investigative site to determine if there are any outlying sites. If an investigative site is unable to enroll a sufficient number of subjects, it may be necessary to combine smaller sites (where enrollment is too low to detect differences) to create "pseudo-sites". Sites with < 7 subjects will be considered eligible for pooling. Eligible sites will be pooled sequentially with other sites in the same region until a sufficient number of subjects are reached (at least seven).

In this process, sites with < 7 subjects will be ordered numerically in ascending order, based on the investigative site number. Sites with < 7 subjects will then be sequentially pooled with other sites in the same region in ascending order by site number until the resulting pseudo-site has at least 7 subjects meeting the PP population definition, thus creating a new "pseudo-site". The process will begin again with the next sequential small site until all small sites have been pooled with other sites in the same region into a pseudo-site. Pooling of the centers will be done prior to study unblinding.

The proportions and 95% confidence intervals for the primary efficacy endpoint data will be presented by pseudo-site to determine if any outlying sites exist.

8.7 Multiple Comparisons/Multiplicity

A hierarchical closed-testing procedure will be utilized for the primary and the secondary efficacy endpoint as noted in Section 7.6.1.1.

8.8 Examination of Subgroups

Subgroup analysis of age (< 65 years, ≥ 65 years), sex (Female, Male), race group (White, Non-white), ethnicity (Hispanic-Latino, not Hispanic-Latino), site geography (outside the US, Eastern

US*, Southern US*, Northern US*, Western US*) and number of previous episodes of CDI recurrence at baseline will be conducted on the primary efficacy analysis if a sufficient sample size exists for each subgroup. The primary efficacy analysis will also be conducted based on vancomycin use for the qualifying CDI episode ≤14 days and > 14 days using subjects who were stratified based on Vancomycin Alone. These analyses will include all analysis populations (mITT and PP) to understand any differences.

*States included in the identified US geographical regions are as follows:

- Eastern US: Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Rhode Island, Vermont, Virginia, West Virginia
- Southern US: Alabama, Arkansas, Georgia, Florida, Kansas, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas
- Northern US: Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Nebraska, North Dakota, South Dakota, Wisconsin
- Western US: Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming

Other sub-groups may be considered if there are enough subjects in each group such as; qualifying CDI event treated as inpatient vs. outpatient, the severity of the qualifying event based on the ATLAS scores, treatment length or type of antibiotics taken at the time of the CDI qualifying event, and others yet undetermined.

Cochran-Mantel-Haenszel tests will be used to assess homogeneity of the odds ratios in the subgroups.

9 COMPUTER SOFTWARE

All analyses will be performed by Medpace using Version 9.3 or later of SAS® software. All summary tables and data listings will be prepared utilizing SAS® software.

10 APPENDIX: BAYESIAN ANALYSIS PLAN

Adaptive Design Report for the Phase 3 Trial (2017-01) for Evaluating the Efficacy and Safety of Rebiotix RBX2660 (microbiota suspension) for the Treatment of Recurrent Clostridium Difficile Infection

Submitted to Rebiotix

August 7, 2019

Introduction

This document describes the adaptive design for a randomized clinical trial to support administration of Rebiotix RBX2660 (microbiota suspension) via enema for the prevention of recurrent *Clostridium difficile* infection (CDI). The trial will enroll and treat up to 270 subjects with recurrent CDI and includes up to two interim analyses at a minimum of 160 and 220 subjects complete (i.e., completed the 8-week follow-up assessment and/or completion of Treatment Outcome form after receiving the blinded treatment). The first interim analysis may occur after 160 subjects complete, pending FDA review of the proposed adaptive design. Early trial success or failure may be declared at either of the interims depending on the strength (either positive or negative) of the observed data. The primary analysis of the trial will be a Bayesian hierarchical model, which formally incorporates data from a previous randomized Phase 2b study (Protocol 2014-01) of RBX2660.

Treatment Arms

Subjects will be randomized to one of the following treatments:

- Treatment: A single dosing of RBX2660 via enema
- Control: A single dosing of placebo via enema

A 2:1 randomization ratio will be used to incentivize patient enrollment in the trial.

Primary Analysis Population

The primary analysis will be conducted using the Modified Intent-to-Treat (mITT) population described in the statistical analysis plan (SAP). Subjects in the mITT population who withdraw due to CDI related symptoms will be counted as treatment failures. Up to 270 patients will be treated in the trial. Patients in the mITT population who dropout prior to completion will be treated as failures; the effect of dropout is described in the section **Operating Characteristics Under Dropout**.

Primary Endpoint

The primary endpoint of the trial is recurrence of CDI within 8 weeks of blinded treatment. Treatment success is defined as the absence of CDI diarrhea through 8 weeks after completing the blinded study treatment.

Statistical Model

Historical Data

The primary statistical analysis will use a hierarchical model to dynamically borrow information about the treatment effect from the previous Phase 2b trial (Protocol 2014-01). The efficacy objectives of the previous study were to assess the effectiveness of two dosing strategies of RBX2660 in patients with CDI. In one dosing group, patients received two enemas of RBX2660 (administered 7± 2 days apart). In the alternate dosing group, patients received a single enema of RBX2660 along with a placebo (administered 7± 2 days apart). Both treatments were compared to patients in a control arm who received 2 administrations of placebo via enema (administered 7± 2 days apart). Data from this trial are shown in Table 1.

Table 1: *Data from Phase 2b trial (Protocol 2014-01).*

Group	Dosing	# Subjects	# Responders	Response Rate
A	2 doses of RBX2660	41	25	0.610
B	Placebo	44	19	0.432
C	1 dose of RBX2660	42	25	0.595

The hierarchical model will incorporate data from groups B and C, excluding group A due to the difference in treatment dosage.

The patients enrolled in this study are expected to be enrolled from a population similar to the Phase 2b trial. In the case that subjects from the Phase 2b trial differ substantially from the current study, the proposed borrowing strategy allows for heterogeneity between study populations and dynamically borrows less information when the data from the new study differs from the Phase 2b study.

For example, one difference between the studies is that patients with a single recurrence of CDI may be enrolled in this trial whereas they were excluded from the Phase 2b trial. We expect that this change will minimally impact the baseline rate of CDI recurrence observed in the new study. However, if we did observe a control response rate that differed substantially from the Phase 2b study, the design would borrow less information, reducing the role of the historical data in our inference. This effect is described in the section **Strength of the Borrowing**.

Hierarchical Model

The data from the Phase 3 trial and Phase 2b trial will be analyzed using a Bayesian hierarchical model. The analysis will be conducted using a modified intent to treat population. Let $N_{k,s}$ be the number of subjects assigned to treatment k (k = T for patients who receive a single dose of RBX2660 and k = C for patients who receive the control) in study s ($s = 1$ for the current Phase 3 study, $s = 2$ for the previous Phase 2 study). We model the number of responders, $X_{k,s}$, in each arm/study as

$$X_{k,s} \sim \text{Binomial}(N_{k,s}, p_{k,s})$$

where $p_{k,s}$ is the underlying event rate for arm k in study s. The event rates are transformed to the log-odds scale and modeled as:

$$\log\left(\frac{p_{C,s}}{1-p_{C,s}}\right) = \alpha_s$$

for the control arms and

$$\log\left(\frac{p_{T,s}}{1-p_{T,s}}\right) = \alpha_s + \theta_s$$

for the treatment arms. The parameter θ_s represents the effect of RBX2660, relative to placebo, on the log-odds scale for trial s. Hierarchical models are used to borrow information about the treatment and control effects across studies. The following prior is used for the control rates across the two trials:

$$\alpha_s \sim N(\alpha, \tau_\alpha^2) \quad \text{for } s = 1, 2$$

$$\alpha \sim N(0, 10^2)$$

$$\tau_\alpha^2 \sim \text{Inverse Gamma}(0.001, 0.1)$$

A similar prior is used for the treatment effects across the two trials:

$$\theta_s \sim N(\theta, \tau_\theta^2) \quad \text{for } s = 1, 2$$

$$\theta \sim N(0, 10^2)$$

$$\tau_\theta^2 \sim \text{Inverse Gamma}(0.01, 0.01)$$

The priors on α and θ are chosen to be conservative and emphasize a prior assumption of a control response rate near 0.5 with a treatment effect centered around 0. The priors on the hierarchical variance terms τ_α^2 and τ_η^2 are parameterized by their location μ and weight ω (rather than the traditional $shape = \omega/2$ and $scale = \mu^2 \cdot \omega/2$ parameterization). These priors were chosen to encourage dynamic borrowing between the studies. This distribution is defined by the density:

$$f(x|\mu, \omega) \propto \frac{e^{-\mu^2\omega/2x}}{x^{\omega/2+1}}$$

The selection of the borrowing hyper parameters is discussed in the **Strength of the Borrowing** Section of this design report.

Success Criteria

Our primary interest lies in the treatment effect (TE) for the current Phase 3 study (s = 1):

$$TE = p_{T,1} - p_{C,1}$$

The primary goal of the trial is to demonstrate the efficacy of RBX2660 by testing the hypothesis:

$$H_0: TE \leq 0 \qquad \text{vs.} \qquad H_A: TE > 0$$

We test this hypothesis by calculating the posterior probability of superiority, Pr($TE > 0$ | Data) , which is equivalent to Pr($\theta_1 > 0$ | Data). The thresholds for defining success at each analysis are summarized in Table 2.

Table 2: *Success criteria at each analysis. The * indicates that the success criteria may be updated at the final analysis to ensure the correct amount of α −spend.*

Analysis	Criteria	Subjects complete
Interim 1	Pr($TE > 0$ | Data) > 0.99943	Minimum 160
Interim 2	Pr($TE > 0$ | Data) > 0.99943	220
Final	Pr($TE > 0$ | Data) > 0.99943*	Maximum 270
Final	Pr($TE > 0$ | Data) > 0.97706	Maximum 270

The success criteria of 0.99943 was chosen using a Pocock spending function for analysis performed at 160, 220, and 270 patients complete with a cumulative α spend of 0.00125. Success at this level is considered appropriate for approval based on results from a single trial.

Uncertainty from several factors make assessing the exact α −spend at each analysis difficult to quantify. First, the definition of the mITT population allows for the final analysis to be performed with fewer than 270 complete subjects. In addition, the first interim analysis may be performed when more than 160 subjects are enrolled and treated. Finally, the second interim analysis may be omitted if the first interim analysis occurs sufficiently close to 220 subjects enrolled. Therefore, to account for this uncertainty, the success criteria at the final analysis will be updated to reflect the information spend at each interim analysis and ensure a cumulative α spend of 0.00125. This change will be made based on the fraction of information at each interim analysis and does not depend on the outcome data for the trial.

The lower, secondary, threshold of 0.97706 at the final analysis allows for success at the less demanding $\alpha = 0.025$ level. The second threshold at the final analysis controls the overall Type I error rate for the trial at 2.5% (without borrowing), accounting for the α spend at the interims and higher look at the final analysis.

Futility Stopping

The trial will stop for futility at either interim if the predictive probability (PP) of success at the final analysis is sufficiently small. The predictive probabilities will be computed using a final threshold of 0.97706 and assuming a final enrollment of 270 subjects. Bayesian predictive distributions incorporate the uncertainty around the model parameters as well as the uncertainty in the outcomes of subjects not yet enrolled. The trial will stop for futility if there is limited probability that the trial will reach the success criterion, even at the maximum sample size. The criteria for declaring futility at each interim are shown in Table 3.

Table 3: *Futility criteria at each analysis.*

Analysis	Criteria	Subjects complete
Interim 1	PP(Success) < 0.01	Minimum 160
Interim 2	PP(Success) < 0.01	220

Strength of the Borrowing

The hyper-priors on τ_α^2 and τ_θ^2 are chosen to encourage dynamic borrowing from the Phase 2b trial. Dynamic borrowing is the idea that when the observed data in the current trial is similar to the previous trial, the strength of our conclusions should increase. On the other hand, if the newly observed data is quite different from the previous data, the prior data should contribute less information to our inference.

Consider two extremes. The first is the case of full borrowing, where data from the previous study is added to data from the new trial and inference is performed using the full weight of both studies. Alternatively, we could do no borrowing, ignoring information from the previous trial and conducting our analysis with only the new information. The hyper-priors on τ_α^2 and τ_θ^2 are specified to provide dynamic borrowing, behaving similarly to the full borrowing approach when the observed data resembles the Phase 2b data and mimicking the no borrowing approach when the results differ in the two trials. The effect of the dynamic borrowing is illustrated in Figure 1.

Figure 1: *Effective sample size (ESS) borrowed on the treatment arm for different observed treatment and response rates at the first interim.*

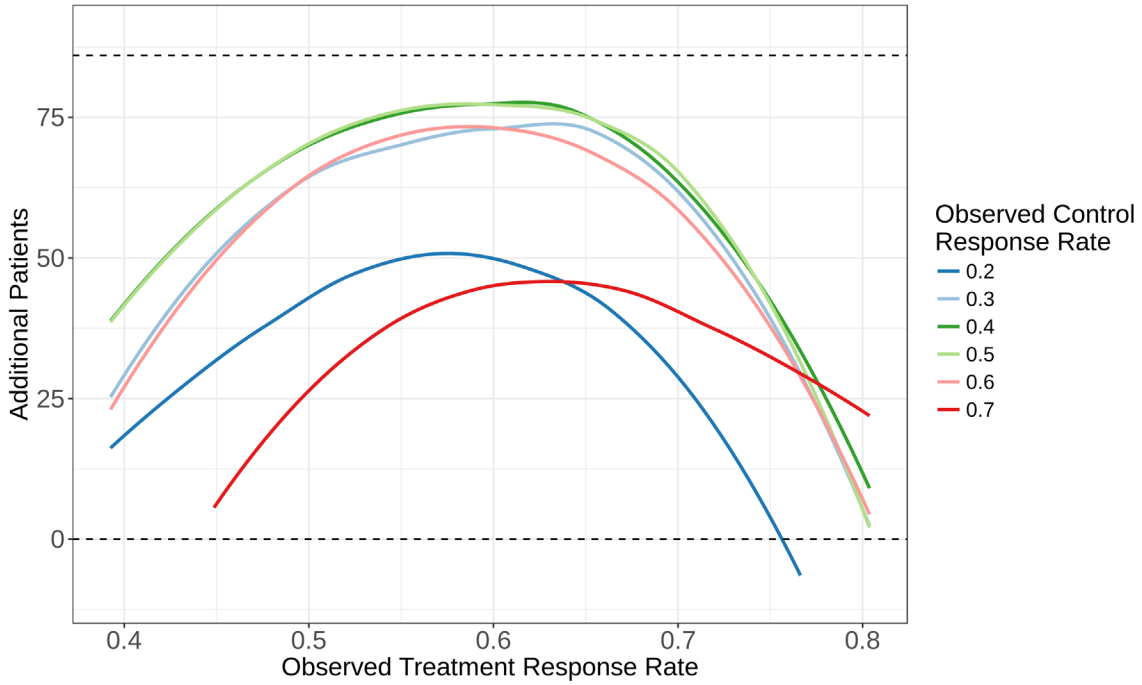


Figure 1 summarizes the amount of borrowing from the hierarchical model for different observed treatment success proportions ($X_{T,1}/N_{T,1}$, the x-axis) and different observed control proportions ($X_{C,1}/N_{C,1}$, differently colored lines). At the first interim analysis, 104 subjects are assumed to be enrolled to the treatment arm with 56 subjects enrolled in the control arm. The y-axis shows the effective number of additional treatment patients borrowed by the model, given by the equation

$$\text{Additional patients} = \frac{\text{Var}(\theta_{full})}{\text{Var}(\theta_1)}(N_1 + N_2) - N_2$$

The quantity $\text{Var}(\theta_1)$ is the posterior variance of the treatment effect under the hierarchical model and $\text{Var}(\theta_{\text{full}})$ is the variance of the treatment effect under the full borrowing model. The ratio of these terms is the relative effective sample size, which measures the uncertainty in the hierarchical model relative to the full model. When this ratio is small, there is substantially more uncertainty in the hierarchical model compared to the full model.

The horizontal dashed lines correspond to 0 (no borrowing from the Phase 2b data) and 86 (full borrowing of the Phase 2b data of Groups B and C). The hierarchical model borrows the most patients from the Phase 2b trial when the observed treatment response rate is close to the treatment response rate in the previous study (a rate of 0.595). As the observed control rate differs from the historical rate, fewer additional patients from the previous trial are used and the amount of borrowing diminishes. When the observed control rate differs from the historical rate, borrowing is strongest when the log-odds of the treatment effect are similar to the phase 2b data. The chosen hyper-priors create appropriate dynamic behavior, decreasing uncertainty when then observed data is in concordance with the previous trial and increasing uncertainty when the Phase 3 trial results differ from the Phase 2b trial.

Example Trials

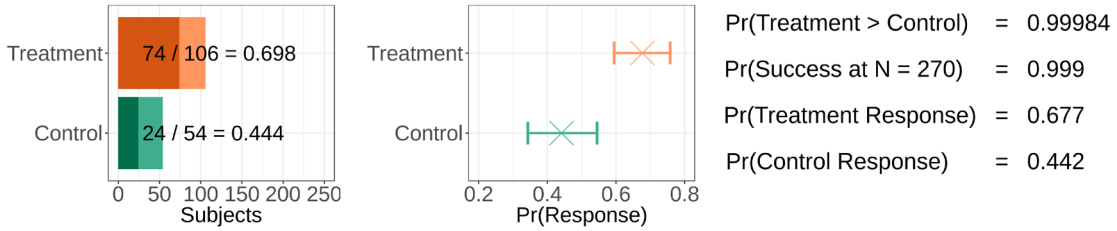
This section contains several example trials selected to highlight aspects of the analysis and design. For each example trial result, we show the following information at each analysis.

- **Observed data**: a bar plot showing the number of patients with complete data enrolled per arm (total width of the bar) as well as the number of responders (darker shaded region).
- **Credible intervals**: plot showing the fitted response rate for each arm (the X) and 95% credible interval for each arm.
- **Pr(Treatment > Control)**: the posterior probability of a positive treatment effect in the Phase 3 trial, used for declaring trial success according to Table 2.
- **Pr(Success at N = 270)**: the predictive probability of success at the final analysis. The trial stops for futility when this probability drops below 0.001.
- **Pr(Treatment Response):** posterior mean of the treatment arm response probability
- **Pr(Control Response):** posterior mean of the control arm response probability

Example Trial 1

The first interim occurs when the minimum of N=160 subjects have complete data (i.e., completed their 8-week follow-up assessment and/or completion of Treatment Outcome form after receiving the blinded treatment). Example 1.1 shows the status of the trial at the first interim.

Example 1.1: *Status of the trial at the 1st interim*

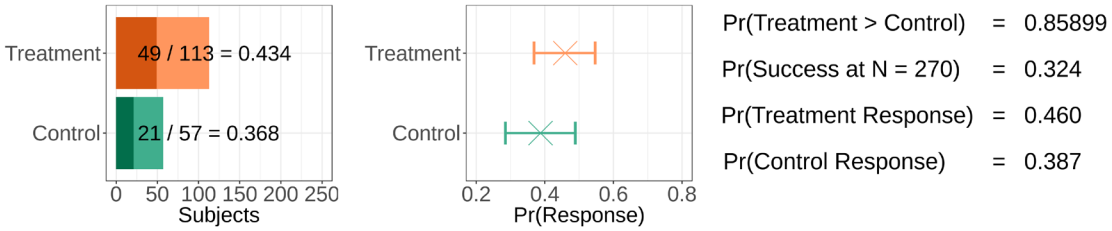


The response rate on the control arm is 44.4%, which is similar to the Phase 2b data, and the response rate on the treatment arm is 69.8%, which is larger than observed in the Phase 2 trial. The posterior estimates of the treatment response rate are smaller than the observed rate due to the borrowing from the Phase 2b data. The posterior probability of a positive treatment effect exceeds 0.99943 and success is declared.

Example Trial 2

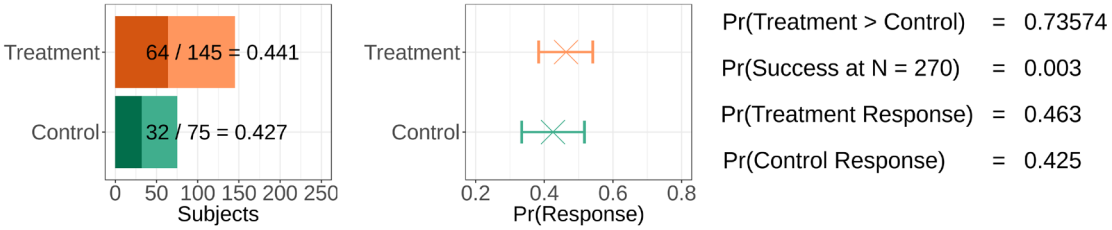
The first interim occurs when N=170 subjects have complete data. Example 2.1 shows the status of the trial at the first interim.

Example 2.1: *Status of the trial at the 1st interim*



The observed response rates in this scenario are much smaller than in the Phase 2b data, with an observed treatment effect of 0.066. While the predictive probability of success is small (0.324), it is sufficient to continue the trial to the second interim.

Example 2.2: *Status of the trial at the 2nd interim*

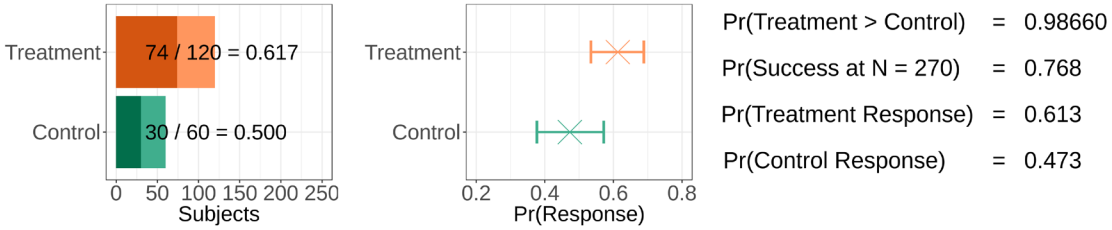


The second interim, shown in Example 2.2, occurs when 220 patients have complete data. The observed response rates have increased in both arms, though the observed treatment effect is smaller (0.014). While the estimated treatment effect is positive, the probability of declaring success at the final analysis is 0.003 which is sufficiently small to stop the trial for futility.

Example Trial 3

In this example, the first interim analysis occurs when 180 subjects have been enrolled. The observed control rate is 0.5, somewhat higher than seen in the Phase 2b trial. Example 3.1 shows the status of the trial at the first interim

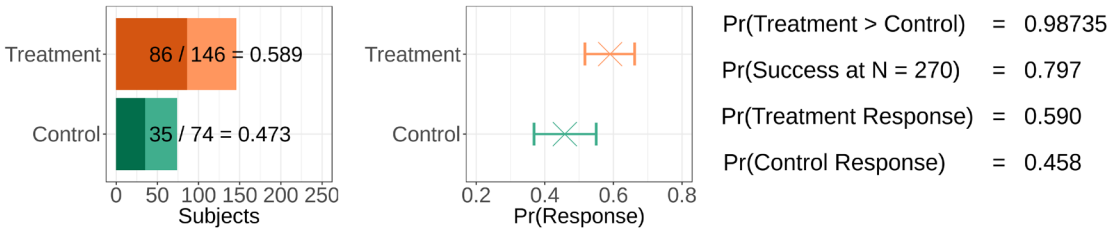
Example 3.1: *Status of the trial at the 1st interim*



The posterior estimate of the control response rate (0.473) is somewhat lower than

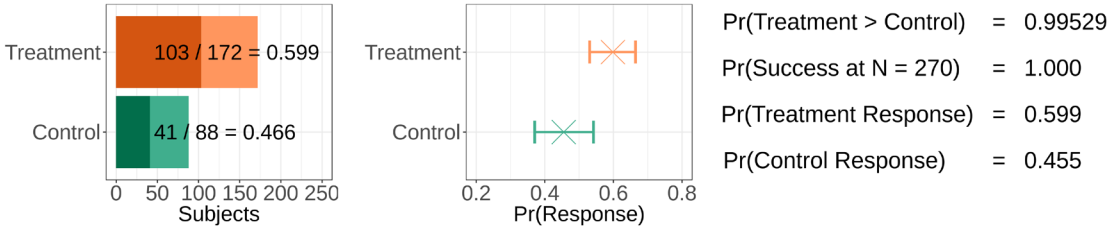
observed due to the borrowing and the posterior probability of a treatment effect is insufficiently high to declare success at the first interim.

Example 3.2: *Status of the trial at the 2nd interim*



At the second interim (N=220, shown in Example 3.2), the observed response rates in both arms have been reduced. The trial has not shown sufficient evidence to declare success or futility and continues enrollment, though the predictive probability of success at the final analysis remains optimistic (0.797).

Example 3.3: *Status of the trial at the final analysis*



The final analysis is conducted when 270 patients have completed follow-up and 260 subjects qualify for the mITT population (Example 3.3). The single trial success criteria is updated to 0.99931, using the information fraction and 0.99943 criteria at the first two interims (0.692 and 0.846) to ensure an overall α −spend of 0.00125. The trial is declared a success at the 0.025 level, though it does not meet the single trial success criteria.

Performance Characteristics

The performance characteristics of the design were determined through trial simulation. We assume several scenarios for the underlying response probability for the control and treatment arms, considering control probabilities in {0.2, 0.3, 0.4, 0.45, 0.5, 0.6, 0.7} and additive treatment effects of {0.0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3}. The scenario with a control probability of 0.45 and a treatment effect of 0.15 corresponds most closely to what was observed during the Phase 2b trial. The scenarios with treatment effects of 0 correspond to null scenarios. For the purposes of simulation, we assume that patients are randomized in 3 patient blocks, that interim analysis were performed with exactly 160 and 220 subjects completed, and that the 270 mITT subjects have complete data at the final analysis.

Operating Characteristics

We simulated 1,000,000 virtual trials for each scenario. The results of the simulations are summarized using the following operating characteristics.

- **Pr(Success):** the cumulative probability of declaring success at the analysis. For example, Pr(Success) under the section Interim 2 includes the probability of declaring success at either Interim 1 or Interim 2. For the final analysis, Pr(Success) (0.99943) and Pr(Success) (0.97706) include the cumulative probability of declaring success at the single trial and lower success levels respectively.
- **Pr(Futility):** the cumulative probability of declaring futility at the analysis
- **# Subjects**: the average number of subjects enrolled in the trial
- **Bias**: the average error in estimating the true (simulated) treatment effect.

Table 4 shows an example set of operating characteristics when the simulated control response probability is 0.45.

Table 4: *Operating characteristics of the design when the underlying control response probability is 0.45.*

	Interim 1		Interim 2		Final analysis	
Effect	**Pr(Success)**	**Pr(Futility)**	**Pr(Success)**	**Pr(Futility)**	**Pr(Success) (0.99943)**	**Pr(Success) (0.97706)**
0	0.001	0.445	0.001	0.732	0.002	0.042
0.05	0.006	0.204	0.012	0.424	0.018	0.204
0.1	0.037	0.065	0.072	0.158	0.106	0.528
0.15	0.141	0.014	0.257	0.036	0.354	0.833
0.2	0.363	0.002	0.564	0.005	0.695	0.969
0.25	0.651	0.000	0.838	0.000	0.919	0.997
0.3	0.868	0.000	0.965	0.000	0.989	1.000

Under the null scenario where both the control and treatment arms have a 45% true response rate, 45% of trials stop for futility at the first interim analysis (N = 160) and 0.1% of trials meet the success criterion. By the second interim, 73% of trials have stopped for futility. Assuming a final threshold of 0.97706, the probability of trial success at or before the final analysis is 0.042. The type I error of this trial is inflated beyond the nominal 2.5% level because of the borrowing from the positive phase 2 data.

In the scenario with a treatment effect of 0.15 (which most closely reflects the observed data from Phase 2), the trial has 83% power. Broken down by analysis, 14% of trials declare success at the first interim, another 12% of trials declare success at the second interim, 9% of trials reach the higher success threshold of 0.99875 at the final analysis and 48% of trials declare success at the lower hurdle at the final analysis. When the true

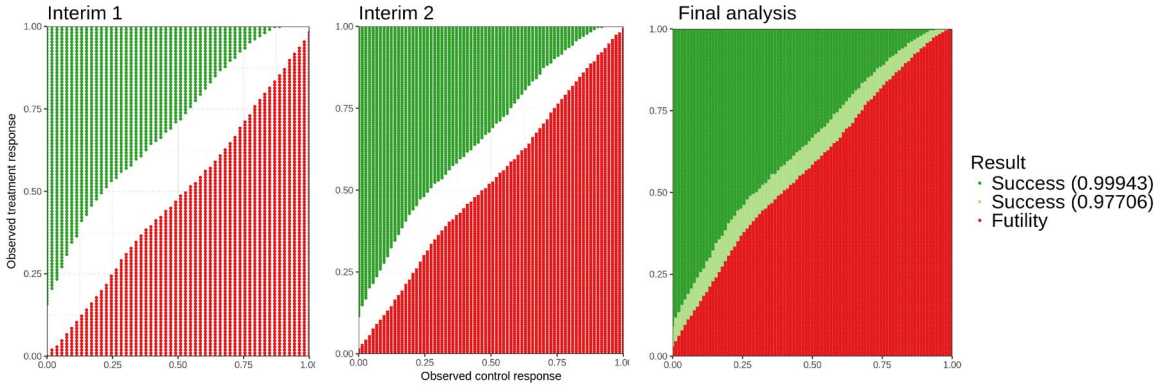
effect is 0.10, the overall power drops to 52% with 77% of trials progressing to the final analysis.

Operating characteristics for simulations with control response probabilities in {0.2, 0.3, 0.4, 0.5, 0.6, 0.7} are included at the end of this section.

Decision Rules

The decision rules of the trial at each analysis can be characterized in terms of the observed effect sizes (# responders / # subjects enrolled) in each arm. The effects leading to either futility or success, colored by result, are shown in in Figure 2.

Figure 2: *Scatterplots of the observed response rates for both arms, colored by result.*



The design has a clear decision boundary at the first interim, declaring futility or success when the observed treatment effect is sufficiently large or small. The gap between success and futility narrows at the 2nd interim, as more information about the arms informs decision making. The effect size required to win is smallest at the final interim.

Operating Characteristics with No Borrowing

For comparison we include the operating characteristics of an alternate design without borrowing. This design uses the same interim schedule as the borrowing design but does not include stopping for futility. Table 6 compares the cumulative probability of declaring success for at each interim for both designs when the simulated control response rate is 0.45.

Table 6: *Operating characteristics for the borrowing design (borrow) and alternate design (none) at each interim analysis when the simulated response rate is 0.45.*

	Interim 1		Interim 2		Final (0.99943)		Final (0.97706)	
Effect	Borrow	None	Borrow	None	Borrow	None	Borrow	None
0	0.001	0.000	0.001	0.001	0.002	0.001	0.042	0.022
0.05	0.006	0.004	0.012	0.007	0.018	0.010	0.204	0.106
0.1	0.037	0.019	0.072	0.038	0.106	0.059	0.528	0.320
0.15	0.141	0.072	0.257	0.141	0.354	0.212	0.833	0.629
0.2	0.363	0.202	0.564	0.362	0.695	0.496	0.969	0.877
0.25	0.651	0.429	0.838	0.655	0.919	0.794	0.997	0.977
0.3	0.868	0.693	0.965	0.883	0.989	0.955	1.000	0.998

For a simulated treatment effect of 0.15, the borrowing design has 20% more power than the alternate design. Broken down by interim, the borrowing design is 7% more likely to declare success at the first interim, 12% more likely to declare success by the second interim, and 14% more likely to reach the higher threshold by the final analysis. The borrowing design is more likely to declare success when the simulated treatment effect is 0, with a 4.2% chance of declaring success compared to 2.2% for the alternate design.

Operating Characteristics Under Drop Out

Per the protocol, subjects who drop out of the mITT population prior to the final analysis for CDI related symptoms will be treated as treatment failures (regardless of arm). Therefore, simulated dropout has the practical effect of reducing the simulated control and treatment rates and the tables in the following second can be used to understand the performance of the design under a range of dropout scenarios.

For example, suppose that the true response rate on the control arm was 0.45, the true treatment effect was 0.20 and the dropout rate was 10%. Then, this trial would have the same operating characteristics of a trial with a simulated 0.405 control rate, a 0.145 treatment effect, and no dropout. The operating characteristics in this scenario can be adequately approximated using Table 8.

Additional Operating Characteristics

This section contains operating characteristics for control response probabilities in *{*0.2, 0.3, 0.4, 0.5, 0.6, 0.7*}*. A description of these tables can be found under the operating characteristics section.

Table 6: *Operating when the underlying control response probability is 0.2*

	Interim 1		Interim 2		Final analysis	
Effect	Pr(Success)	Pr(Futility)	Pr(Success)	Pr(Futility)	Pr(Success) (0.99943)	Pr(Success) (0.97706)
0	0.000	0.462	0.001	0.732	0.001	0.050
0.05	0.003	0.225	0.009	0.449	0.015	0.203
0.1	0.015	0.087	0.044	0.208	0.077	0.466
0.15	0.054	0.026	0.146	0.070	0.241	0.740
0.2	0.153	0.005	0.353	0.015	0.515	0.919
0.25	0.346	0.001	0.629	0.002	0.790	0.986
0.3	0.608	0.000	0.862	0.000	0.949	0.999

Table 7: *Operating when the underlying control response probability is 0.3*

	Interim 1		Interim 2		Final analysis	
Effect	Pr(Success)	Pr(Futility)	Pr(Success)	Pr(Futility)	Pr(Success) (0.99943)	Pr(Success) (0.97706)
0	0.000	0.539	0.000	0.811	0.001	0.025
0.05	0.002	0.301	0.004	0.571	0.007	0.114
0.1	0.010	0.122	0.026	0.285	0.044	0.347
0.15	0.049	0.033	0.119	0.089	0.189	0.679
0.2	0.172	0.006	0.353	0.016	0.495	0.912
0.25	0.419	0.001	0.674	0.002	0.812	0.988
0.3	0.710	0.000	0.901	0.000	0.964	0.999

Table 8: *Operating when the underlying control response probability is 0.4*

	Interim 1		Interim 2		Final analysis	
Effect	Pr(Success)	Pr(Futility)	Pr(Success)	Pr(Futility)	Pr(Success) (0.99943)	Pr(Success) (0.97706)
0	0.000	0.506	0.001	0.788	0.001	0.027
0.05	0.003	0.246	0.007	0.495	0.011	0.152
0.1	0.023	0.084	0.050	0.204	0.076	0.454
0.15	0.103	0.019	0.199	0.051	0.286	0.784
0.2	0.294	0.003	0.489	0.007	0.628	0.955
0.25	0.575	0.000	0.790	0.001	0.891	0.995
0.3	0.827	0.000	0.951	0.000	0.984	1.000

Table 9: *Operating when the underlying control response probability is 0.5*

	Interim 1		Interim 2		Final analysis	
Effect	Pr(Success)	Pr(Futility)	Pr(Success)	Pr(Futility)	Pr(Success) (0.99943)	Pr(Success) (0.97706)
0	0.001	0.389	0.002	0.668	0.003	0.063
0.05	0.010	0.166	0.019	0.358	0.027	0.260
0.1	0.054	0.049	0.101	0.123	0.143	0.595
0.15	0.187	0.009	0.316	0.025	0.418	0.869
0.2	0.431	0.001	0.626	0.003	0.748	0.978
0.25	0.700	0.000	0.870	0.000	0.938	0.998
0.3	0.885	0.000	0.971	0.000	0.992	1.000

Table 10: *Operating when the underlying control response probability is 0.6*

	Interim 1		Interim 2		Final analysis	
Effect	Pr(Success)	Pr(Futility)	Pr(Success)	Pr(Futility)	Pr(Success) (0.99943)	Pr(Success) (0.97706)
0	0.004	0.296	0.006	0.559	0.007	0.108
0.05	0.022	0.117	0.038	0.269	0.052	0.351
0.1	0.089	0.035	0.153	0.089	0.205	0.671
0.15	0.241	0.007	0.384	0.019	0.490	0.894
0.2	0.475	0.001	0.670	0.003	0.789	0.981
0.25	0.733	0.000	0.894	0.000	0.957	0.999
0.3	0.925	0.000	0.986	0.000	0.998	1.000

Table 11: *Operating when the underlying control response probability is 0.7*

	Interim 1		Interim 2		Final analysis	
Effect	Pr(Success)	Pr(Futility)	Pr(Success)	Pr(Futility)	Pr(Success) (0.99943)	Pr(Success) (0.97706)
0	0.004	0.290	0.007	0.555	0.009	0.108
0.05	0.021	0.120	0.036	0.279	0.049	0.325
0.1	0.082	0.033	0.141	0.086	0.197	0.661
0.15	0.246	0.005	0.396	0.013	0.529	0.919
0.2	0.559	0.000	0.760	0.001	0.880	0.995
0.25	0.877	0.000	0.977	0.000	0.995	1.000
0.3	1.000	0.000	1.000	0.000	1.000	1.000

Computational Details

Simulations were run using R (R Core Team 2017) version 3.4.3. Bayesian computations were performed using the Stan (Stan Development Team 2018) version 2.17.2.

Many samples from the posterior distribution are required to determine whether the posterior probability is greater than 0.99943 with sufficient precision. For the power calculations, for each trial we ran a single MCMC sampler for 10,000 iterations. When executing the trial, if the posterior probability of a positive treatment effect is close to the boundary we recommend using more samples to obtain sufficient precision. For example, using 1,000 independent Markov chains with 10,000 post burn-in iterations each.