Statistical Analysis Plan

NCT Number: NCT04759833

Title: Phase 3, Multicenter, Randomized Study with 2 Different Doses of Prucalopride Administered to Male and Female Pediatric Subjects Aged 6 Months to 17 years with Functional Constipation, Consisting of a 12-week Double-blind, Placebo-controlled Part (Part A) to Evaluate Efficacy and Safety Followed by a 36-week Double-blind Extension Part (Part B) to Document Long-term Safety up to week 48

Study Number: TAK-555-3010

Document Version and Date: Version 5.0, 24 November 2023

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

TAKEDA PHARMACEUTICALS

STATISTICAL ANALYSIS PLAN

TAK-555, prucalopride succinate
PHASE 3

Phase 3, Multicenter, Randomized Study with 2 Different Doses of Prucalopride Administered to Male and Female Pediatric Subjects Aged 6 Months to 17 years with Functional Constipation, Consisting of a 12-week Double-blind, Placebo-controlled Part (Part A) to Evaluate Efficacy and Safety Followed by a 36-week Double-blind Extension Part (Part B) to Document Long-term Safety up to week 48

Protocol Identifier	TAK-555-3010
Study Sponsor(s)	Takeda Development Center Americas, Inc. (TDC Americas) 95 Hayden Avenue Lexington, MA 02421 617-349-0200
Authors	[redacted] and [redacted]
Protocol	Amendment 3 27 Oct 2022
SAP Version #	5.0
SAP Date	24 Nov 2023
Status	Final

REVISION HISTORY

Version No:	Issue Date	Summary of Changes	Author
1.0	14 Oct 2019	Final	[redacted]
2.0	22 Feb 2021	Amended to incorporate changes made in protocol version 2.0.	[redacted]
3.0	07 Nov 2022	Amended to incorporate changes due to clarification on calculation of compliance rate, small sites pooling requested by FDA and subject enrollment in Part A and Part B respectively.	[redacted]
4.0	24 May 2023	1. Correct analysis window. 2. Update statistical method for imputation as "predictive mean matching" and update SAS code. 3. Clarify that BM within 24 hours after disimpaction period will not be considered as spontaneous. 4. Clarify details for the calculation of conditional power and add a backup plan imputation in case that one arm has too few subjects by IA1. 5. Add lower limits for Blood Pressure Levels.	[redacted]
5.0	24 Nov 2023	1. Update secondary endpoints to describe subgroups by reporter status as per protocol. 2. Combine Toilet-trained and Non-toilet-trained populations. Perform safety analyses and summarize non-toilet-trained subjects together with those for toilet-trained subjects. 3. Make analysis decision based on study termination due to futility. 4.Update lab ranges for safety endpoints.	[redacted]

SIGNATURE PAGE

Protocol Title: Phase 3, Multicenter, Randomized Study with 2 Different Doses of Prucalopride Administered to Male and Female Pediatric Subjects Aged 6 Months to 17 years with Functional Constipation, Consisting of a 12-week Double-blind, Placebo-controlled Part (Part A) to Evaluate Efficacy and Safety Followed by a 36-week Double-blind Extension Part (Part B) to Document Long-term Safety up to week 48

Protocol Number: TAK-555-3010

Protocol Version: 3.0

Protocol Date: 27 Oct 2022

SAP Version: 5.0

Study Sponsor(s): Takeda Development Center Americas, Inc. (TDC Americas)
95 Hayden Avenue
Lexington, MA 02421

Prepared by: [redacted]
Statistics
Statistics and Quantitative Science, Takeda

DocuSigned by: [redacted]
25-Nov-2023 | 03:23:16 JST

Signature Date

Approved by: [redacted]
Statistics
Statistics and Quantitative Science, Takeda

DocuSigned by: [redacted]
24-Nov-2023 | 20:54:37 CET

Signature Date

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

AE	adverse event
BM	bowel movement
BSFS	Bristol Stool Form Scale (BSFS)
CGI-S	caregiver global impression of severity
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
CRO	contract research organization
CSR	clinical study report
DMC	data monitoring committee
ECG	electrocardiogram
eCRF	electronic case report form
FDA	Food and Drug Administration
FGID	functional gastrointestinal disorders
FoTA	final on treatment assessment
GI	gastrointestinal
IA	interim analysis
ICE	intercurrent event
IRT	interactive response technology
ITT	intention-to-treat
IP	investigational product
LOV	last observed value
LS	least square
LVOT	last value on treatment
MAR	missing at random
MCID	minimum clinically important difference
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
MMRM	mixed model for repeated measures
MNAR	missing not at random
NRS	numerical response scale
NTTIIT	non-toilet trained intent-to-treat
NTTmITT	non-toilet trained modified intent-to-treat
NTTSAF	non-toilet trained safety analysis set

PCS	potentially clinically significant
PedsQL	pediatric quality of life
PGI-S	patient global impression of severity
PK	pharmacokinetic
PP	per-protocol
PT	preferred term
REML	restricted maximum likelihood
QoL	quality of life
SAP	statistical analysis plan
SBM	Spontaneous bowel movement
SD	standard deviation
SRM	standardized response mean
SOC	system organ class
TEAE	treatment-emergent adverse event
WHO	World Health Organization
WOCF	worst observation carried forward

1. INTRODUCTION

This statistical analysis plan (SAP) provides a technical and detailed description of the statistical analyses of efficacy and safety data for Study TAK-555-3010.

Details of the statistical methods stated in the protocol (Version 3.0, Amendment 3, dated 27 Oct 2022) are provided and, where applicable, significant changes from the protocol-specified analyses will be documented, prior to database lock, through a SAP amendment. If additional analyses are required to supplement the planned analyses described in this SAP after the database lock, they may be completed and will be described in the clinical study report (CSR).

2. OBJECTIVES, ESTIMANDS, AND ENDPOINTS

2.1 Study Objectives

2.1.1 Primary Objective

- To evaluate the efficacy of prucalopride during the 12-week double-blind, placebo-controlled treatment phase in toilet-trained subjects with functional constipation who are at least 3 years of age.
- To evaluate the long-term (48 weeks) safety and tolerability of prucalopride in toilet-trained subjects with functional constipation who are at least 3 years of age.

2.1.2 Key Secondary Objectives

- To evaluate the efficacy of prucalopride on signs and symptoms of functional constipation (stool consistency, straining, and stool frequency responder index), during the 12-week double-blind, placebo-controlled treatment phase in toilet-trained subjects with functional constipation who are at least 3 years of age.

2.1.3 Other Secondary Objectives

- To evaluate the effect of prucalopride on the frequency of fecal incontinence.
- To evaluate safety and tolerability of prucalopride over 12 weeks of treatment in toilet trained subjects who are at least 3 years of age.
- To evaluate the pharmacokinetics (PK) of prucalopride by combining sparse PK blood sampling that will be conducted during the 12-week double-blind, placebo-controlled phase together with the data from previous pediatric studies for further population PK analysis (separate reporting).

2.1.4 Exploratory Objectives

- To further explore the efficacy of prucalopride on signs and symptoms of functional constipation during the 12-week double-blind, placebo-controlled treatment phase in toilet-trained subjects with functional constipation who are at least 3 years of age.
- To explore the efficacy of prucalopride during the 12-week double-blind, placebo-controlled treatment phase in non-toilet-trained subjects with functional constipation who are at least 6 months of age.
- To evaluate the long-term (48 weeks) safety and tolerability of prucalopride in non-toilet-trained subjects with functional constipation who are at least 6 months of age.
- To evaluate safety and tolerability of prucalopride over 12 and 48 weeks of treatment in non-toilet-trained subjects with functional constipation who are at least 6 months of age.

- To explore the relationship between PK and efficacy/safety endpoints of interest.

2.2 Estimands

The alignment of the study objectives, study design, data collection, study conduct, data analysis and data interpretation are contained within the framework of the estimands, which is described in this section. The estimand framework for the primary objective consists of several components, ie, the estimand type, the population, variable/endpoint, intercurrent events, and population summary.

The primary estimand is a composite type estimand which accounts for treatment discontinuation due to lack of efficacy as an intercurrent event.

The estimand population is the modified intent-to-treat (ITT) set, which includes toilet-trained randomized subjects who are at least 3 years of age.

The primary endpoint is the average change from baseline in number of spontaneous bowel movements (SBMs) per week derived from the (diary) data over 12 weeks, collected during the placebo-controlled part (Part A). Any bowel movement (BM) that occurs within 24 hours after intake of the rescue medication or within 24 hours after disimpaction period will not be considered as spontaneous.

Intercurrent events (ICEs) with influence on the estimand are identified as follows:

- Discontinuation of treatment due of lack of efficacy and/or adverse event (AE). The "composite strategy" will be utilized to address the ICEs where discontinuation due to lack of efficacy or AE will be considered as undesired treatment outcome in the analysis.
- Usage of rescue medication.
- This ICE is captured through the endpoint definition. Data will continue to be collected for subjects on rescue medication.

The population summary of the estimand framework includes missing data, primary analysis methodology, and assessment of robustness through sensitivity analyses.

Missing data will be imputed using a hybrid imputation approach prior to analyses of primary and key secondary efficacy endpoints. Subjects who discontinued due to lack of efficacy or AE during the study will be considered as having undesired treatment outcome in the primary analysis. Subsequently, their missing data will be imputed using the worst observation carried forward (WOCF) under the missing not at random (MNAR) assumption. All other discontinuation/intermittent missingness will be imputed using multiple imputation under the missing at random (MAR) assumption. For multiple imputation, the weekly number of SBMs

will be imputed by treatment group using a multivariate stepwise approach using a fully conditional specification regression method with predictive mean matching method due to the non-negative nature of the imputed values, which is similar to the regression method except that it imputes a value randomly from a set of observed values whose predicted values are closest to the predicted value for the missing value from the simulated regression model (Heitjan and Little 1991; Schenker and Taylor 1996). A minimum threshold of zero for the weeks with missing values will also be set. The number of k observations to be used for the mean matching will be set to 5. Missing baseline weekly number of SBMs, if any, will be imputed using age and baseline global severity score. Missing post-baseline weekly number of SBMs will be imputed using all previous weeks in a stepwise fashion. Twenty multiply imputed datasets will be generated. Lastly, the imputed data under MNAR and MAR will be combined to obtain a complete dataset containing all subjects in the analysis population and will be used to perform the pre-specified statistical modeling for primary and key secondary efficacy endpoints using mixed effects model for repeated measures (MMRM) for continuous endpoints and Cochran Mantel-Haenzel (CMH) test for binary endpoint.

The primary estimand will be estimated from a MMRM based on restricted maximum likelihood (REML) using the imputed data from the hybrid imputation approach. The MMRM will include treatment group, age groups, study week, treatment-group-by-study-week interaction as fixed effects, baseline number of SBM/week as a covariate and subject as a random effect. An unstructured variance-covariance matrix will be used to model the within-subject errors for both treatment groups. The average change from baseline over 12 weeks will be estimated least square (LS) means by the above MMRM. The treatment difference in LS means between active treatment group versus placebo will be estimated. P-value, treatment difference in LS means and associated 95% confidence interval (CI) from the multiple imputed datasets will be combined using Rubin's rules, as implemented in the PROC MIANALYZE procedure (Rubin 1987).

Several sensitivity analyses, see Section 6.1.1, will be executed to demonstrate the robustness of the primary analysis.

2.3 Study Endpoints

2.3.1 Primary Endpoint

The primary endpoint (in toilet-trained subjects who are at least 3 years of age) is as follows:

- The average change from baseline in number of SBMs[1] per week derived from the (diary) data over 12 weeks, collected during the placebo-controlled part (Part A).

2.3.2 Key Secondary Endpoints

- Efficacy endpoints in toilet-trained subjects who are at least 3 years of age:
  - The average change from baseline in stool consistency (based on Bristol Stool Form Scale [BSFS] score), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase. Summaries will be reported separately for subjects aged <8 years and 8 to 17 years based on reporter status (parent/child). The average will be calculated as sum of all scores divided by the number of evaluable days in a 7-day period.
  - The average change from baseline in straining (based on a 3-point Likert scale, assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase. The average will be calculated as sum of all scores divided by the number of evaluable days in a 7-day period.
  - The proportion of responders with a responder defined as a subject having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of the 12 weeks of placebo-controlled part (Part A), including 3 of the last 4 weeks.

2.3.3 Other Secondary Endpoint

- Efficacy endpoints in toilet-trained subjects who are at least 3 years of age:
  - Proportion of subjects with fecal incontinence per week during the 12-week treatment period.

2.3.4 Exploratory Endpoints

- Efficacy endpoints in toilet-trained subjects who are at least 3 years of age:
  - The average change in worst abdominal pain score over the past 24 hours (based on a Wong-Baker faces scale in subjects <8 years and the 11-point Numerical Response Scale

[1] A BM is defined as spontaneous (SBM) if not preceded within a period of 24 hours by the intake of rescue medication or within 24 hours after disimpaction period.

[NRS] in subjects ≥8 years), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase.

- Proportion of subjects with an average of ≥3 SBMs per week and increase of ≥1 SBM compared to baseline during a 12-week double-blind, placebo-controlled treatment phase.
- Proportions of subjects with an average of ≤1 SBMs per week during the 12-week double-blind, placebo-controlled treatment phase.
- Proportion of subjects with an average of ≥1 day(s) with rescue medication intake per week.
- Proportion of subjects assessed with retentive posturing at monthly visits during the 12-week treatment period.
- Retentive posturing is defined as the attempt to preserve continence by vigorous contraction of the gluteal muscles. Children with retentive posturing will be typically tight legged, tiptoed, and/or will have a back-arching posture.
- Global evaluation and quality of life (QoL) health economics and outcomes research endpoints: Patient Global Impression of Severity (PGI-S), Caregiver Global Impression of Severity (CGI-S), and Pediatric Quality of Life Inventory™ (PedsQL) Gas and Bloating Domain Gastrointestinal (GI) Symptoms Scale (Placebo-controlled Part [Part A]):
  - Absolute values and change from baseline.
  - For PedsQL Gas and Bloating Domain GI Symptoms Scale Standardized response mean (SRM): classified as no, small, moderate, and large effect.
- Proportion of subjects with ≤2 signs/symptoms from the Rome IV Criteria following the 12-week double-blind, placebo-controlled treatment phase

– Efficacy endpoints in non-toilet-trained subjects who are at least 6 months of age:

- Proportion of subjects with an average of ≥3 SBMs per week and increase of ≥1 SBM compared to baseline during the 12-week double-blind, placebo-controlled treatment phase.
- Number of SBMs per week in categories ≤1 and >1 during the 12-week double-blind, placebo-controlled treatment phase.
- The average change from baseline in stool consistency (based on BSFS score), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase.

2.3.5 Safety Endpoints

– The proportion of subjects with treatment-emergent adverse events (TEAEs) (serious, non-serious, related, non-related) and the proportion of subjects with clinically relevant laboratory test abnormalities, electrocardiogram (ECG) findings, vital signs findings, or new findings in physical examination over 12 and 48 weeks of treatment.

2.3.6 PK Endpoint

– For all subjects who consent, regardless of toilet-training and age, sparse PK blood sampling will be conducted during the 12-week double-blind, placebo-controlled phase. The obtained plasma concentrations, together with the data from previous pediatric studies (PRU-USA-12, PRU-USA-24 and SHP555-303), will be used for further population PK analysis (see protocol Section 9.8 Pharmacokinetic and Pharmacokinetic/Pharmacodynamic Analyses) and the refined model will be used to characterize prucalopride's PK properties, and the relationship between PK and efficacy/safety endpoints of interest may be explored (separate reporting). Individual plasma concentrations with corresponding sampling times post-dose will be tabulated per visit. A graphical presentation will be used to visualize the plasma concentration-time relationship. Descriptive statistics will be calculated per visit and at specific post dose time intervals.

3. STUDY DESIGN

3.1 Overall Design

This is a Phase 3, multicenter, randomized study consisting of a 12-week double-blind, placebo-controlled Part (Part A) followed by a 36-week double-blind safety extension part (Part B) to document safety and tolerability up to 48 weeks in male and female pediatric and adolescent subjects with functional constipation as defined by the modified Rome IV criteria for child/adolescent Functional Gastrointestinal Disorders (FGID).

The study will start with a 10- to 33-day screening period, including a disimpaction for all subjects, a 12-week double-blind placebo-controlled part (Part A) followed by a 36-week long-term double-blind (for dose) safety extension Part (Part B), and a follow-up call approximately 4 weeks after the last administration of the investigational product (IP).

3.2 Randomization

Randomization will be implemented by interactive response technology (IRT) and individual subject treatment is automatically assigned by the IRT. Subject randomization, by means of the IRT, occurs at 2 time points, at the beginning of study parts A and B as described below.

- Placebo-controlled Part (Part A):
- The IRT system will ensure an approximately 1:1:1 balance between the 3 treatment groups (prucalopride low dose, prucalopride high dose, and placebo) in each stratum.
- Randomization will be performed with the permuted blocks and will be stratified by toilet-trained status, age group (<12 years, 12 to 17 years) for those toilet-trained, and average number of SBM/week (≤1; >1) during the screening period. The randomization block size(s) will remain unknown to the investigator sites as well as the study team.
- Safety Extension Part (Part B):
- The IRT system will ensure an approximately 1:1 balance between the 2 treatment groups (prucalopride low dose and prucalopride high dose).
- Re-randomization of the placebo treatment group at the end of the placebo-controlled Part A will be performed with the permuted blocks and will not be stratified. The randomization block size will remain unknown to the investigator sites as well as the study team.

Drug accountability will be collected using IRT and completed in the site drug accountability logs.

3.3 Blinding

Matching placebo solution/tablets will be available such that the blinding can be assured in identical packaging.

The subject, investigator, study coordinator, sponsor, and the entire study processing team will remain blinded to the treatment assignment. The set-up of the randomization system will ensure that the blind is maintained and will not reveal treatment allocation to any unauthorized personnel.

The interim analysis (IA) will be conducted by an unblinded statistics/programming team at the contract research organization (CRO) and this team is independent from the study team in that the unblinded team will have no other study oversight or inputs into analysis planning. To ensure study integrity, the unblinded IA results will be reviewed by an external data monitoring committee (DMC) who will make recommendations to the study sponsor.

3.4 Sample Size and Power Consideration (Toilet-trained Subjects)

There is limited literature available with information of efficacy measures in children with functional constipation. For that reason, the data from a previous pediatric study with prucalopride (SPD555-C303) was used to estimate current primary endpoint - change from baseline in the average number of SBMs per week derived from the (diary) data over 12 weeks. In that 8-week study only the last 4 weeks (Weeks 5-8) of the treatment phase were used for evaluation of primary efficacy[2] and over this period the average change in SBMs/week was approximately 1.40 SBM/week for placebo and 1.52 SBM/week for prucalopride. The overall average change in SBMs/week was 1.5 with a pooled standard deviation (SD) of 2.30 (SD in placebo arm was 2.02 and in prucalopride treatment arm 2.57). This change is based on a selection of subjects with <3 SBM/week at baseline, corresponding to the inclusion criteria of the current study.

An attempt was made to get an estimate of the minimum clinically important difference (MCID) in average change from baseline of number of SBM/week, based on the data of the previous pediatric study with prucalopride (SPD555-C303).

For this, the change from baseline in number of SBM/week was estimated per outcome score of the global evaluation of efficacy of treatment scale (the anchor) after 8 weeks of treatment. This

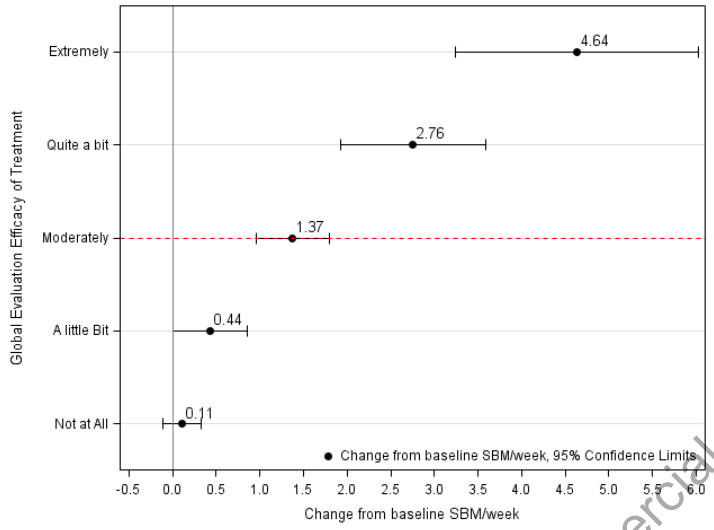
---

[2] In study SPD555-C303, the first two weeks of the double-blind treatment phase were not used for the primary efficacy analysis because of a possible effect from the use of an enema or oral laxative agent during the screening phase to remove an impaction. weeks 3 and 4 were used to evaluate the efficacy for a possible dose adjustment after 4 weeks of treatment. Therefore, the primary endpoint was evaluated over weeks 5 to 8.

scale had 5 outcome scores: not at all -, a little bit -, moderately - quite a bit- and extremely effective).

Figure 1 presents the average change from baseline over all subjects, per outcome of the global evaluation of efficacy of treatment.

Figure 1. Average Change from Baseline in SBMs (Study SPD555-C303)



SBM: spontaneous bowel movement

If we consider a moderate efficacy of treatment as a MCID, then it can be seen that those subjects have an average change from baseline of 1.37 SBM/week.

The 6 key efficacy studies with prucalopride in adults, showed a treatment difference of 1.66 SBM/week for the subgroup of patients with <3 SBM/week at baseline.

Based on all this information and since this is a pediatric study, a difference of 1.40 SBM/week between a prucalopride treatment arm and placebo was chosen as a (clinically meaningful) treatment effect for the sample size estimation.

As we expect to observe a higher treatment effect with one of the doses of prucalopride in current study, as compared to the previous pediatric study SPD555-C303, this might result in more variability and therefore a slightly higher value for the pooled SD of the change from baseline will be assumed to be 2.50.

The sample size was estimated through statistical simulations based on the Hochberg step-up procedure to control the type I error rate for primary efficacy endpoint. These simulations showed that with 80 toilet-trained subjects who are at least 3 years of age per treatment arm,

Part A of the study will have at least 90% power to detect a treatment difference of 1.40 in primary efficacy endpoint between at least one active dose versus placebo assuming pooled SD of 2.5, using a two-sided two-sample t-test at a significance level of 5% based on the Hochberg step-up procedure to control the type I error rate for primary efficacy endpoint.

In addition, for exploratory purposes, it is targeted to enroll a maximum of 15 non-toilet-trained subjects.

4. STATISTICAL ANALYSIS SETS

4.1 Statistical Analysis Sets (Toilet-trained subjects who are at least 3 years of age)

4.1.1 Screened Set

The screened set will consist of all subjects who have signed an informed consent form.

4.1.2 Enrolled Set

The enrolled set will consist of all subjects who have signed informed consent and also either passed inclusion/exclusion criteria or were randomized.

4.1.3 ITT Analysis Set

The ITT analysis set will consist of all randomized subjects.

4.1.4 Modified Intent-to-treat Analysis Set

The modified intent-to-treat (mITT) analysis set is applicable to Study Part A and will consist of all randomized subjects who receive at least 1 dose of IP. The treatment arm will be the arm to which they were randomized, regardless of what treatment they received.

4.1.5 Safety Analysis Set

For Study Part A, the safety analysis set will consist of randomized subjects receiving at least 1 dose of IP from Study Part A. Subjects will be analyzed according to treatment received.

For Study Part B, the safety analysis set will consist of all randomized who receive at least 1 dose of IP from Study Part B. Subjects will be analyzed according to the treatment received.

4.1.6 Per-protocol Analysis Set

The per-protocol (PP) analysis set is applicable to Study Part A and will consist of all toilet-trained randomized subjects who are at least 3 years of age and who do not have major protocol deviations that may affect the primary efficacy endpoint. These major protocol deviations will be finalized and documented prior to study unblinding.

Examples of such deviations are provided below:

- Violations of inclusion and/or exclusion criteria: A subject who violated any inclusion and/or exclusion criteria will be excluded from the PP Analyses Set.
- Compliance with IP: A subject, who was less than 80% or more than 120% compliant (can occur when the subject/caregiver intentionally or unintentionally overdoses) with his/her assigned treatment, as assessed by tablet count or liquid volume from the date of first dispensing to the date of last dose, will be excluded.

- Incorrect timing of assessments: A primary/key secondary endpoint assessment performed outside a ±3 day window for the nominal visit time will be classified as a major protocol deviation. Note that this rule does not apply to ediary data, which are collected daily, and endpoint are averages over each week.
- Prohibited concomitant medications that could have a potential effect on efficacy endpoints. Protocol violations based on prohibited concomitant medications will be defined at the at the final data review meeting before database lock.
- Taking wrong IP, ie, not the IP as randomized.
- Based on the deviations as determined at database closure subjects with other major protocol deviations might be excluded from the per-protocol set. All subjects excluded will be listed with the reason for exclusion.

4.1.7 Completers Analysis Set

The Completers Analysis Set is applicable to Study Part A and will consist of all toilet-trained subjects who are at least 3 years of age in the mITT Analysis Set who have an average number of SBM available for all 12 weeks in the Placebo-controlled Part (Part A) of the study. The subject in this set must have at least 4 days with completed diary data in each of the 12 weeks to calculate the average number of SBMs.

4.1.8 PK Analysis Set

PK Analysis Set: The PK Analysis Set is defined as all subjects regardless of age in the Safety Analysis Sets and for whom at least 1 PK sample is evaluable.

5. STUDY SUBJECTS (TOILET-TRAINED SUBJECTS)

5.1 Disposition of Subjects

A listing of all screen failures (ie, subjects who were screened but not randomized) will be presented along with reasons for screen fail; where applicable details of any AEs will be available in the data listings. Subject disposition will be presented separately by study part (A and B).

For Study Part A, subject disposition will be summarized for the Screened Set. The number of subjects who were included in and excluded from each defined analysis set (ie, Screened, Enrolled, ITT, mITT, Safety, and PP Analysis Set) will be summarized by treatment group and overall, except for the Screened Set and the Enrolled set, which will be summarized only overall.

The number and percentage of subjects who completed and prematurely discontinued will be presented for each treatment group and overall, for the Safety Analysis Set. Reasons for premature discontinuation as recorded on the termination page of the electronic case report form (eCRF) will be summarized (number and percentage) by treatment group and overall, for the Safety Analysis Set. All subjects who prematurely discontinued will be listed by discontinuation reason for Enrolled Set.

The duration of enrollment, in days, will be summarized for each site, country, and overall. Duration of enrollment will be calculated as (last date of contact for any subject at that site - the first date of informed consent for any subject at that site +1).

For Study Part B subject disposition will be summarized similarly as above except as explained below. Rather than the screened set the initial set will consist of all subjects that complete Study Part A. The only analysis set to be tallied is the safety analysis set, as all other sets are not applicable to Part B. Subject tallies will also be summarized based on Part A treatment assignment (ie, placebo or prucalopride).

5.2 Demographic and Other Baseline Characteristics

The description below will apply to each Study Part.

Descriptive summaries of demographic and baseline characteristics will be presented by treatment group and overall Safety Analysis Set.

The following demographic characteristics will be summarized: age (years), age (categorical), sex, ethnicity, country/region, race, body weight (kg), height (cm), and body mass index (kg/m^2), removal of fecal impaction (yes/no), duration of constipation and for Study Part B summary

prior treatment assignment in Study Part A. For continuous characteristics, the mean, SD, median, Q1 and Q3, minimum and maximum values will be presented. For categorical characteristics, the number of subjects and percentage will be presented.

5.3 Medical History

Medical history will be collected at the Screening Visit and will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 23.1 or newer. A listing will be provided using the Safety Analysis Set. The description below will be applied to subjects in the Part A safety set separately for each the toilet-trained and non-toilet-trained subjects.

The medical history will be summarized (number of subjects and percentage) by system organ class (SOC) and preferred term (PT) for each treatment group and overall, for the Safety Analysis Set.

5.4 Prior Medication

Prior medication is defined as any medication with the start date and end date prior to the date of the first dose of IP. Prior medications will be coded using the World Health Organization (WHO) Drug Dictionary. The description below will be applied to subjects in the Part A safety analysis set.

Prior medication information must be recorded in the subject's source documents. The prior medication usage will be summarized by the number and proportion of subjects in each treatment group and in overall subjects within each therapeutic class and preferred term for the Safety Analysis Set. Multiple medication usage by a subject in the same category will be counted only once. All prior treatment will be listed for the Safety Analysis Set.

5.5 Concomitant Medication

Concomitant medication will be coded using the WHO Drug Dictionary. The description below will be applied to subjects in the safety analysis set.

Concomitant medication is defined as any medication with a start date prior to the date of the first dose of IP and continuing (ongoing) after the first dose of IP or with a start date between the dates of the first and last doses of IP, inclusive.

Any medication with a start date after the date of the last dose of IP will not be considered a concomitant medication.

The concomitant medication usage will be summarized by the number and proportion of subjects in each treatment group receiving each medication and in overall subjects within each

therapeutic class and preferred term for the Safety Analysis Set. Multiple medication usage by a subject in the same category will be counted only once.

All concomitant medication will be listed for the Safety Analysis Set.

5.6 Exposure to IP

Exposure to IP for the Safety Analysis Set will be summarized in terms of treatment duration, which is calculated as the number of days from the date of first dose of IP taken to the date of the last dose of IP taken, inclusively. Descriptive statistics (n, mean, SD, minimum, median, and maximum) will be presented to describe the exposure to IP by treatment group and overall. The above summary will be produced for each Study Part.

If two sources of the end date are available the date collected on site will be used to determine last dose of IP. See Section 12.5 for missing last dose date.

5.7 Measurements of Treatment Compliance

IP dosing compliance for a specified period is defined as the total number of days that the subject or caregiver report in the e-diary that the medication was taken by a subject during that period divided by the expected number of days medication to be taken in the same time period (difference in days as reported on the eCRF dosing form) multiplied by 100. Descriptive statistics for IP compliance will be presented by treatment group for each period between 2 consecutive visits as well as for the whole double-blind evaluation phase for the study for the Safety Analysis Set. These will include summary statistics (n, mean median, std, min and max) as well as categorization in appropriate categories (eg, <80%, 80-120%, >120%). The above summary will be produced for each Study Part.

5.8 Protocol deviations

Protocol deviations will be captured and classified (eg, critical, major, minor) by the CRO as documented in the Protocol Deviation Management Plan. The Takeda study team will review the protocol deviations and their classification throughout the study and before treatment unblinding and database lock. Decisions of the review will include accuracy of protocol deviations categorization.

Protocol deviations will be summarized by classification and site for each treatment group and overall, for the Safety Analysis Set. Protocol deviations will be listed for the ITT Analysis Set.

6. EFFICACY ANALYSIS (TOILET-TRAINED SUBJECTS)

All efficacy analyses will be based on the mITT for toilet trained subjects who are at least 3 years of age unless otherwise specified. Baseline for efficacy endpoints, that are assessed at study visits, is defined as the last observed value prior to taking the first dose of IP (based on dates). Baseline for efficacy endpoints derived from the daily diary data, is defined as the average value over all diary days observed prior to taking the first dose of IP, excluding the disimpaction period. Baseline for the PedsQL Gas and Bloating Module will be the screening assessment prior to disimpaction. All efficacy analyses will be conducted according to the treatment assigned.

All statistical tests will be 2-sided hypothesis tests performed at the 5% level of significance for main effects. All confidence intervals will be 2-sided 95% confidence intervals, unless stated otherwise. Control of Type I error is discussed in Section 6.4.

6.1 Derivation and Analyses of Primary Efficacy Endpoint

The primary efficacy endpoint is defined as the average change from baseline in number of SBMs per week derived from the (diary) data over 12 weeks.

A BM is defined as spontaneous (SBM) if not preceded within a period of 24 hours by the intake of rescue medication or within 24 hours after disimpaction period.

The null and alternative hypotheses for the primary efficacy analysis are:

- Null hypothesis1 (H10): There is no difference in average change from baseline in number of SBMs/week derived from the (diary) data over 12 weeks between the low dose and placebo.
- Alternative hypothesis1 (H1A): There is a difference in average change from baseline in number of SBMs/week derived from the (diary) data over 12 weeks between the low dose and placebo.
- Null hypothesis2 (H20): There is no difference in average change from baseline in number of SBMs/week derived from the (diary) data over 12 weeks between the high dose and placebo.
- Alternative hypothesis2 (H2A): There is a difference in average change from baseline in number of SBMs/week derived from the (diary) data over 12 weeks between the high dose and placebo.

The primary endpoint will be derived as follows:

1. For each subject, the number of SBMs will be calculated (using the formula below) for each of the 12 weeks in the placebo-controlled part (Part A) of the study and for baseline.

2. The (average) number of SBMs during the screening period (baseline) will be calculated as follows:
   - $Number\ of\ SBM\ (baseline) = \frac{7*(total\ frequency\ of\ SBMs\ during\ the\ screening\ period\ \dagger)}{(number\ of\ days\ with\ observation\ during\ the\ screening\ period\ \dagger)}$
   - † Excluding the days of the disimpaction procedure (3 to 12 days).
3. For the calculation of weekly number of SBM, it is required that there are at least 4 days with completed diary data (ie, evaluable days), otherwise the weekly number of SBM will be set to missing. A day is considered evaluable (day with diary information) if at least some information on BM was recorded (ie, 'no BM today' or the time of at least 1 BM was recorded).
4. The (average) number of SBM per week will be calculated as follows, where i is a study week within 1 to 12, inclusively:
   - $Number\ of\ SBMs\ (Week\ i) = \frac{7*(total\ frequency\ of\ SBMs\ in\ Week\ i)}{(number\ of\ days\ with\ observation\ in\ Week\ i)}$
5. The (average) change from baseline in number of SBM/week can be derived for all weeks during Placebo-controlled Part of the study.
6. Data from Step 5 will be used in an MMRM model for the estimation and analysis of the primary endpoint (average change in SBM/week over 12 weeks).

Note that BMs during the disimpaction period(s), which can take 3 to 12 days during the screening period and BMs within 24 hours after rescue medication intake during Placebo-controlled Part of the study, will not be counted.

Missing data will be imputed using a hybrid imputation approach prior to the analysis of the primary efficacy endpoint. The primary efficacy endpoint will be analyzed using a MMRM. The MMRM will include treatment group, age groups, study week, treatment-group-by-study-week interaction as fixed effects, baseline number of SBM/week as a covariate and subject as a random effect. An unstructured variance-covariance matrix will be used to model the within-subject errors for both treatment groups. The average change from baseline over 12 weeks will be estimated (LS means) by the above MMRM. The treatment difference in LS means between active treatment group versus placebo will be estimated. P-value for treatment difference in LS means and associated 95% CI from the multiple imputed datasets will be combined using Rubin's rules, as implemented in the PROC MIANALYZE procedure. Both unadjusted and Hochberg's adjusted p-values will be presented. For multiplicity adjustment and decision rules, see Section 6.4. The restricted maximum likelihood method will be used, with an unstructured (UN) covariance structure. If the UN covariance structure fails to converge, as an alternative, the Compound symmetry covariance matrix will be applied. The primary endpoint will be summarized per treatment group and evaluated over the mITT analysis set.

Sample SAS codes are provided in Appendix 2.

6.1.1 Missing Data and Sensitivity Analyses of Primary Efficacy Endpoint

The primary method for missing data handling will be based on a hybrid imputation approach under a composite estimand, with details shown in Section 2.2.

Missing Diary Days

For the calculation of weekly number of SBM, it is required that there are at least 4 days with completed diary data (ie, evaluable days), otherwise the weekly number of SBM will be set to missing. In the case when there are at least 4 days with diary data, the (average) number of SBM per week will be calculated as follows, where i is a study week within 1 to 12, inclusively:

$$\text{Number of SBMs (week } i) = 7 * \frac{total\ frequency\ of\ SBMs\ in\ week\ i}{number\ of\ days\ with\ observation\ in\ week\ i}$$

Sensitivity Analyses

- Sensitivity Analyses: The following sensitivity analysis for missing data handling will be conducted to examine the robustness of the primary efficacy analysis results:
  - On treatment multiple imputation: For any subject with missing values, the estimates for the missing weekly frequencies will be based on the treatment group of that subject (ie, on-treatment multiple imputation). In other words, for a prucalopride subject, data from the placebo group will not be used to impute missing weekly frequencies. This is the difference between this method and the placebo multiple imputation method (see below). Twenty imputations will be performed.
  - Placebo multiple imputation: For any subject with missing values, the estimates for the missing weekly frequencies will be based on data of the placebo group. In other words, for a prucalopride subject, data from the placebo arm will be used to impute the missing weekly frequencies. Twenty imputations will be performed.
  - Completer analysis: This analysis will include only subjects who have an average number of SBMs available for all 12 weeks in the Placebo-controlled Part (Part A) of the study (the completers analysis set). See Section 4.1.7 for more details.
  - PP analysis: This analysis will include only subjects from the PP analysis set.
  - MMRM analysis with additional covariate and factors: The proposed MMRM model for primary efficacy analysis, including additional factors (center, tablet vs oral solution, global severity score) and covariate age. Note: When age is analyzed as a covariate, age group will not be included as a factor in the MMRM; for purpose of convergence, pooling of smaller centers might be needed. Centers with fewer than 5 subjects in the Full

Analysis Set will be pooled into 1 of 4 regional pools. The pooling strategy will be reassessed and documented prior to unblinding to ensure that the regional pools are not over representative, ie, contain more subjects than the large enrolling centers, wherein large regional pools will be divided with centers being selected at random.

- Tipping-point analysis: A tipping-point multiple imputation analysis will be conducted to assess the robustness of the missing data handling for primary efficacy endpoint if the primary analyses conclude statistical significance. The tipping point is defined as the difference in the average number of SBM/week between a treatment arm and placebo arm at which the study conclusion (ie, the statistical significance) changes. To find this tipping point a systematic shift will be applied to the imputed values (for weeks with missing number of SBM/week) in each of the TAK-555 treatment arms respectively, assuming that the tipping point that reverses the study conclusion is between –3 and 0 with a step size of -0.2 (or adapt as appropriate). The multiple imputation under MNAR assumption will be performed by searching for the tipping point that reverses the study conclusion.

A tipping point analysis will be performed as described below.

The steps to be used for tipping point analysis are as follows:

Step 1: Specify the shift parameter = 0.0, -0.2, -0.4, -0.6,…,-3.0 (0 to -3 with a step size of -0.2, or adapt as appropriate)

Step 2: Impute the missing average number of SBM/week (for weeks with missing number of SBM/week) on TAK-555 treatment arms using PROC MI under the MNAR assumption with the shift parameter to be applied to the TAK-555 arms only. Generate n=20 imputed datasets for each of the shift parameters.

Step 3: Perform the same MMRM analysis model as specified for the primary endpoint analysis for each imputed dataset.

Step 4: Obtain the point estimate of treatment difference, 95% CI and p-value using Rubin's combination rules via PROC MIANALYZE procedure.

Step 5: The tipping point can be identified while the result is no longer statistically significant (ie, p-value >0.05).

The mITT analysis set for subjects who are at least three years of age and toilet-trained will be used for all sensitivity analyses except for the completer and PP analyses.

Sample SAS codes for these sensitivity analyses are provided in Appendix 2.

6.2 Analyses of Key Secondary Efficacy Endpoints

Key secondary efficacy endpoints are:

- Efficacy endpoints in toilet-trained subjects who are at least 3 years of age:
  - The average change from baseline in stool consistency (based on BSFS score), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase. Summaries will be reported separately for subjects aged <8 years and 8 to 17 years based on reporter status (parent/child). The average will be calculated as sum of all scores divided by the number of evaluable days in a 7-day period.
  - The average change from baseline in straining (based on a 3-point Likert scale), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase. The average will be calculated as sum of all scores divided by the number of evaluable days in a 7-day period.
  - The proportion of responders with a responder defined as a subject having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of the 12 weeks of placebo-controlled part (Part A), including 3 of the last 4 weeks

The null and alternative hypotheses for the key secondary efficacy analysis are:

- Null hypothesis 3 (H30): There is no difference in average change from baseline in stool consistency (based on BSFS score) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the low dose and placebo.
- Alternative hypothesis 3 (H3A): There is a difference in average change from baseline in stool consistency (based on BSFS score) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the low dose and placebo.
- Null hypothesis 4 (H40): There is no difference in average change from baseline in stool consistency (based on BSFS score) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the high dose and placebo.
- Alternative hypothesis 4 (H4A): There is a difference in average change from baseline in stool consistency (based on BSFS score) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the high dose and placebo.
- Null hypothesis 5 (H50): There is no difference in average change from baseline in straining (based on a 3-point Likert scale) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the low dose and placebo.
- Alternative hypothesis 5 (H5A): There is a difference in average change from baseline in straining (based on a 3-point Likert scale) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the low dose and placebo.

- Null hypothesis 6 (H60): There is no difference in average change from baseline in straining (based on a 3-point Likert scale) assessed as the weekly average over 12 weeks during the 12-week double-blind, placebo-controlled treatment phase between the high dose and placebo.
- Alternative hypothesis 6 (H6A): There is a difference in average change from baseline in straining (based on a 3-point Likert scale) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase between the high dose and placebo.
- Null hypothesis 7 (H70): There is no difference in proportion of responders having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of 12 weeks between the low dose and placebo.
- Alternative hypothesis 7 (H7A): There is a difference in proportion of responders having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of 12 weeks between the low dose and placebo.
- Null hypothesis 8 (H80): There is no difference in proportion of responders having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of 12 weeks between the high dose and placebo.
- Alternative hypothesis 8 (H8A): There is a difference in proportion of responders having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of 12 weeks between the high dose and placebo.

The key secondary efficacy endpoints will be analyzed as follows:

Missing data will be imputed using a hybrid imputation approach prior to analyses of key secondary efficacy endpoints.

The continuous endpoints using MMRM with the same covariates and factors as for the primary endpoint.

- Similar to the primary endpoint analysis, the average change from baseline in stool consistency (based on BSFS score), and in straining (based on a 3-point Likert scale) assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase, will be analyzed using MMRM. The MMRM will include treatment group, age groups, study week, treatment-group-by-study-week interaction as fixed effects, baseline number of SBM/week as a covariate and subject as a random effect.
- For the proportion of responders, treatment arms will be compared using CMH test, controlling for the stratification variables age and baseline number of SBMs/week. An overall combined p-value, response rates and difference in response rates plus 95% CI will be derived using Rubin's rules, as implemented in the PROC MIANALYZE procedure. The

CMH test statistic will be normalized using the Wilson-Hilferty transformation before combining.

– The overall CMH test statistic and p-value (which are based on a chi-square distributed statistic) will be derived using the Wilson-Hilferty transformation (to normalize a chi-square distributed statistic).

The methodology for all derivation is described in a PharmaSUG 2013 paper (SP03): Combining Analysis Results from Multiply Imputed Categorical Data by Bohdana Ratitch, Ilya Lipkovich and Michael O'Kelly.

For multiplicity adjustment and decision rules, see Section 6.4.

The secondary endpoint will be summarized per treatment group and evaluated over the mITT analysis set.

6.3 Analyses of Other Secondary Efficacy Endpoints

– Other secondary efficacy endpoints in toilet-trained subjects who are at least 3 years of age:
  - Proportion of subjects with fecal incontinence per week during the 12-week treatment Period.

The other efficacy endpoints will be analyzed as follows:

For the other efficacy endpoints based on diary data, the weekly number of SBMs scores will be derived in a same way as for the primary endpoint. Missing data will be imputed using a hybrid imputation approach prior to analyses of key secondary efficacy endpoints.

The other secondary and exploratory efficacy endpoints will be summarized descriptively by treatment arm and evaluated over the mITT analysis set for those subjects who are at least 3 years of age and toilet-trained.

The treatment arms will be compared per week (Weeks 1 to 12) using CMH test, controlling for the stratification variables age and baseline number of SBMs/week. An overall combined p-value, response rates and difference in response rates plus 95% CI will be derived using Rubin's rules, as implemented in the PROC MIANALYZE procedure. The CMH test statistic will be normalized using the Wilson-Hilferty transformation before combining.

The methodology for all derivation is described in a PharmaSUG 2013 paper (SP03): Combining Analysis Results from Multiply Imputed Categorical Data by Bohdana Ratitch, Ilya Lipkovich and Michael O'Kelly.

No multiplicity adjustment will be performed for the analysis of other secondary or exploratory efficacy endpoints.

6.4 Multiplicity Adjustment and Decision Rules

The global family-wise error rate will be controlled at an alpha of 0.05 for the 8 hypotheses resulting from comparing each prucalopride dose group (low dose and high dose) with placebo for the primary endpoint and the three key secondary endpoints using a Hochberg-step-up procedure [(Dmitrienko et al. 2016); FDA 2017, Guidance for Industry: Multiple Endpoints in Clinical Trials. US Department of Health and Human Services, FDA, CDER, CBER].

This approach will use the Hochberg step-up procedure for the primary endpoint dose-comparisons vs. placebo and the truncated Hochberg step-up procedure for each of the three key secondary endpoint dose comparisons vs. placebo, in a stepwise manner in the order below:

- Endpoint P [primary efficacy endpoint is defined as the average change from baseline in number of SBMs per week derived from the (diary) data over 12 weeks].
- Endpoint S1 [The average change from baseline in stool consistency (based on BSFS score), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase]
- Endpoint S2 [The average change from baseline in straining (based on a 3-point Likert scale), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase]
- Endpoint S3 [The proportion of responders with a responder defined as a subject having an increase of ≥1 SBM/week compared to baseline and ≥3 SBMs/week for at least 9 out of the 12 weeks of placebo-controlled part (Part A), including 3 of the last 4 weeks]

The resulting eight null hypotheses of no treatment effect are denoted by H10 through H80 and are grouped into four families corresponding to the four endpoints and families are tested sequentially beginning with first one:

Family1 (F1) = {H10, H20} (Endpoint P)

Family2 (F2) = {H30, H40} (Endpoint S1)

Family3 (F3) = {H50, H60} ((Endpoint S2)

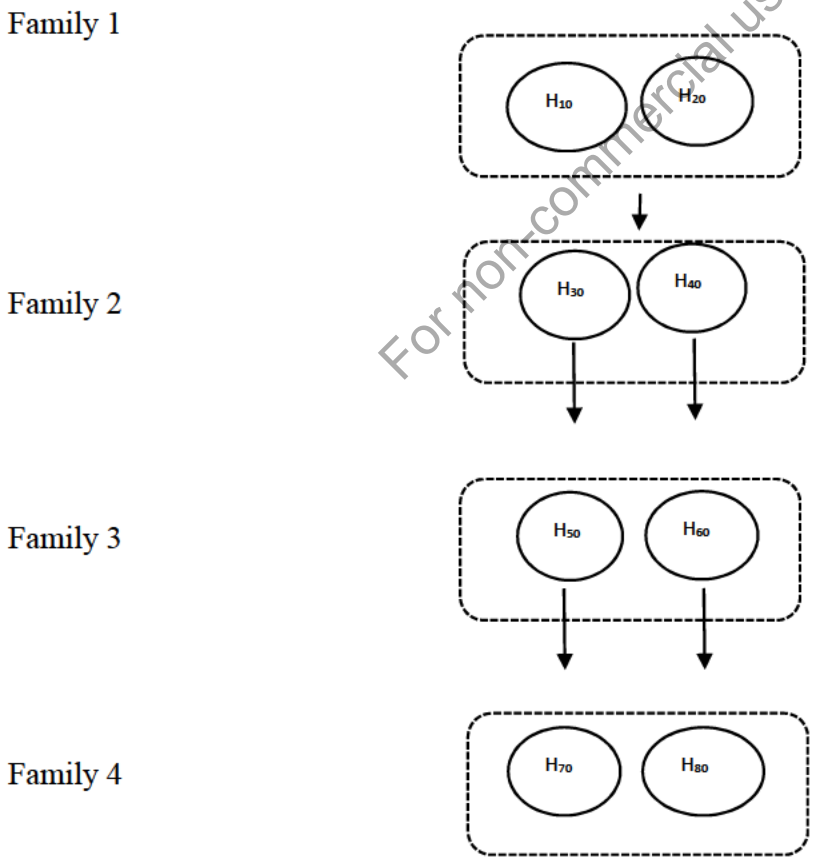
Family4 (F4) = {H70, H80} (Endpoint S3)

Here Family 1 includes the null hypotheses related to the primary endpoint (H10 and H20 correspond to the comparison between Low Dose vs Placebo, and High Dose vs Placebo, respectively.

Families 2,3, and 4 include the null hypotheses related to the key secondary endpoints (H30, H50, and H70 correspond to the comparison between Low Dose vs Placebo; H40, H60, and H80 correspond to the comparison between High Dose vs Placebo).

The eight null hypotheses are organized into two sequences and serial logical restrictions are imposed within each sequence to account for the clinically meaningful relationships. As shown in the testing diagram in Figure 2, a null hypothesis should become testable only if all preceding hypotheses in the sequence are rejected and non-testable hypothesis should be automatically accepted. For example, the null hypothesis H70 should become testable only if H10, H20, H30, and H50 are all rejected.

Figure 2. Flowchart of Testing Primary Endpoint and Key Secondary Endpoints Sequentially.



The testing algorithm is based on regular Hochberg-step-up procedure for F1 and truncated Hochberg procedure with pre-specified truncation parameters $\gamma1= \gamma2=0.8$ for F2 and F3 respectively. The first key secondary endpoint will be tested only if both comparisons for high dose vs. placebo and low dose vs. placebo for primary efficacy endpoint have demonstrated statistical significance.

The testing algorithm is described as follows:

Step 1: The null hypotheses H10 and H20 are tested using the regular Hochberg at $\alpha=0.05$
Let p(1), p(2) denote the ordered p-values corresponding to the null hypotheses H10 and H20 in the Family 1, where $p(1) < p(2)$.

The regular Hochberg procedure uses the following decision rules:

- Both null hypotheses in Family 1 are rejected if $p(2) < \alpha=0.05$
- Only the null hypothesis corresponding to p(1) is rejected if $p(1) < \alpha/2=0.025$ and $p(2) \geq \alpha=0.05$

Step 2: If both hypotheses H10 and H20 in Family 1 are rejected in Step 1, then the null hypotheses H30 and H40 in Family 2 are tested using truncated Hochberg ($\gamma1=0.8$) at $\alpha=0.05$. Otherwise, hypothesis testing is stopped.

Let p(3), p(4) denote the ordered p-values corresponding to the null hypotheses H30 and H40 in the Family 2, where $p(3) < p(4)$.

The truncated Hochberg procedure uses the following decision rules:

- Both null hypotheses in Family2 are rejected if $p(4) < (1+\gamma1)\alpha/2=(1+0.8)*(0.05/2)=0.045$
- Only the null hypothesis corresponding to p(3) is rejected if $p(3) < \alpha/2 =0.025$, and $p(4) \geq (1+\gamma1)\alpha/2=0.045$

Step 3: The logical restrictions are applied to the null hypotheses H50 and H60 in Family F3, ie, H50 is testable only if H30 is rejected and, similarly, H60 is testable only if H40 is rejected. The overall significance level used in F3 is determined by the number of hypotheses rejected in F2:

- If both null hypotheses H30 and H40 in Family 2 are rejected in F2, Hochberg ($\gamma2=0.8$) is applied to both null hypotheses H50 and H60 in Family F3 at $\alpha=0.05$. To be specific and clear: with $p(5) < p(6)$, reject both H50 and H60 in Family 3 if $p(6)<(1+\gamma2)\alpha/2$; only reject H05 if $p(5)<\alpha/2=0.025$ and $p(6) \geq (1+\gamma2)\alpha/2=0.045$.
- If only one of null hypothesis H30 and H40 is rejected in F2, the corresponding null hypothesis H50 or H60 in F3 is tested using a univariate at $\alpha (1-\gamma1)/2=0.05(1-0.8)/2=0.005$.

Step 4: The logical restrictions are applied to the null hypotheses H70 and H80 in Family F4. The overall significance level used in F4 is determined by the number of null hypotheses rejected in Families F2 and F3:

- If null hypotheses H30, H40, H50 and H60 are rejected in Families F2 and F3, then null hypotheses H70 and H80 in Family 4 are tested using regular Hochberg procedure at $\alpha=0.05$. To be specific and clear: with $p(7) < p(8)$, reject both H70 and H80 in Family 4 if $p(8)<\alpha$; only reject H07 if $p(7)<\alpha/2$ and $p(8) \geq \alpha$.
- If both null hypotheses H30 and H40 are rejected in F2, and only one of null hypothesis H50 and H60 is rejected in F3, the corresponding null hypothesis of either H70 or H80 in family F4 is tested using a univariate test at $\alpha (1- \gamma 2)/2=0.05(1-0.8)/2=0.005$.
- If only one of null hypothesis H30 and H40 is rejected in family F2, and only one of null hypothesis H50 and H60 is rejected in family F3, the corresponding null hypothesis of either H7 or H8 is tested using a univariate test at $\alpha (1-\max(\gamma 1,\gamma 2))/2=0.05(1-0.8))/2=0.005$

6.5 Analyses of Exploratory Efficacy Endpoints

6.5.1 Toilet-trained subjects who are at least 3 years of age

Exploratory efficacy endpoints in toilet-trained subjects who are at least 3 years of age include the following:

- The average change in worst abdominal pain score over the past 24 hours (based on a Wong-Baker faces scale in subjects <8 years and the 11-point NRS in subjects ≥8 years), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase.
- Proportion of subjects with an average of ≥3 SBMs per week and increase of ≥1 SBM compared to baseline during the 12-week double-blind, placebo-controlled treatment phase.
- Proportions of subjects with an average of ≤1 SBMs per week during the 12-week double-blind, placebo-controlled treatment phase.
- Proportion of subjects with an average of ≥1 day(s) with rescue medication intake per week.
- Proportion of subjects assessed with retentive posturing at monthly visits during the 12-week treatment period.
- Retentive posturing is defined as the attempt to preserve continence by vigorous contraction of the gluteal muscles. Children with retentive posturing will be typically tight legged, tiptoed, and/or will have a back-arching posture.
- Proportion of subjects with ≤2 signs/symptoms from the Rome IV Criteria following the 12-week double-blind, placebo-controlled treatment phase

The exploratory efficacy endpoints will be analyzed as follows:

For the exploratory efficacy endpoints based on diary data, the weekly number of SBMs scores will be derived in the same way as for the primary endpoint. Missing data will be imputed using a hybrid imputation approach prior to analyses of key secondary efficacy endpoints.

The other secondary and exploratory efficacy endpoints will be summarized descriptively by treatment arm and evaluated over the mITT analysis set. Where possible data will be presented graphically (over time) to support the interpretation of the results.

- For continuous endpoints, treatment arms will be compared per week (Weeks 1 to 12) using the same MMRM as used for primary endpoint analysis.
- For binary endpoints, treatment arms will be compared per week (Weeks 1 to 12) using CMH test, controlling for the stratification variables age and baseline number of SBMs/week. An overall combined p-value, response rates and difference in response rates plus 95% confidence interval will be derived using Rubin's rules, as implemented in the PROC MIANALYZE procedure. The CMH test statistic will be normalized using the Wilson-Hilferty transformation before combining.

The methodology for all derivation is described in a PharmaSUG 2013 paper (SP03): Combining Analysis Results from Multiply Imputed Categorical Data by Bohdana Ratitch, Ilya Lipkovich and Michael O'Kelly.

No multiplicity adjustment will be performed for the analysis of other and exploratory efficacy endpoints.

6.5.2 Analysis of Efficacy Endpoints for Non-toilet-trained Subjects

Efficacy endpoints in non-toilet-trained subjects who are at least 6 months of age include the following:

- Proportion of subjects with an average of ≥3 SBMs per week and increase of ≥1 SBM compared to baseline during the 12-week double-blind, placebo-controlled treatment phase.
- Number of SBMs per week in categories ≤1 and >1 during the 12-week double-blind, placebo-controlled treatment phase.
- The average change from baseline in stool consistency (based on BSFS score), assessed as the weekly average during the 12-week double-blind, placebo-controlled treatment phase.

The efficacy endpoints in non-toilet-trained subjects will be analyzed as follows:

For exploratory and descriptive purposes, all the efficacy analyses will be descriptive with no formal hypothesis testing. For binary endpoints, frequency and percentage, and for continuous endpoints, descriptive statistics (mean, SD, minimum and maximum) will be presented by each treatment group.

The mITT Analysis Set will be used for efficacy analysis.

6.6 Subgroup Analyses

For explorative and descriptive purposes, the primary and selected secondary endpoints (endpoints based on number of SBMs per week and endpoints based on consistency scoring) will also be summarized and compared between treatment groups within each subgroup as defined by the stratification variables: age group (<12 years, 12 to 17 years) and average number of SBM/week (≤1; >1) during the screening period.

7. SAFETY ANALYSIS (TOILET-TRAINED SUBJECTS AND NON-TOILET TRAINED SUBJECTS)

Safety Analysis of toilet-trained subjects:

The safety analysis will be performed using the Safety Analysis Set. Safety variables include AEs, clinical laboratory variables, vital signs, and ECG variables. For each safety variable, the last value collected before the first dose of IP will be used as baseline for all analyses of that safety variable. Last Value on Treatment (LVOT) will be defined as the last valid assessment obtained after Baseline and whilst on IP. Last Observed Value (LOV) will be defined as the last valid assessment obtained after Baseline.

All safety analyses will be conducted according to the treatment the subject actually received.

All safety data will be summarized for the whole study (Part A and Part B combined), and also separately for Part A and Part B, unless otherwise specified. Part A will be summarized using the placebo, low dose, high dose and Prucalopride overall. Part B will be summarized using the groupings of low dose, high dose and overall. Summaries of Part A and Part B are combined using the groupings placebo, placebo-low dose, placebo-high dose, low dose-low dose, high dose-high dose and Prucalopride overall, unless otherwise specified. Subjects randomized to Placebo in Part A who discontinued from the study prior to starting Part B will also be included in placebo-low dose or placebo-high dose. Subjects that did not reach Part B or did not participate in Part B and received low dose in part A will be included the low/low group. Subjects that did not reach Part B or did not participate in Part B and received high dose in part A will be included the high/high group. If a subject is in the safety set for Part A or Part B they will be included in the combined analyses.

7.1 Adverse Events

An AE (classified by preferred term) that occurs during this study will be considered a TEAE if it has a start date on or after the first dose of double-blind IP or if it has a start date before the date of the first dose of double-blind IP but increases in severity on or after the date of the first dose of double-blind IP. An AE that occurs more than 5 days (5 x half-life of 24 hours) after the date of the last dose of double-blind IP will not be counted as a TEAE. If two sources of the end date of treatment are available, the date collected on site will be used to determine last dose of IP.

A summary of the number of subjects with TEAEs will be presented, including the number and percentage of subjects with any TEAEs, serious TEAEs, TEAEs related to IP and TEAEs leading to discontinuation of IP.

The number and percentage of subjects reporting TEAEs in each treatment group will be tabulated by SOC and preferred term; by SOC, preferred term, and maximum severity. TEAEs considered related to IP will also be summarized by SOC and preferred term. If more than 1 AE occurs with the same preferred term for the same subject, then the subject will be counted only once for that preferred term using the most severe and most related occurrence for the summarization by severity and by relationship to IP.

The incidence of common TEAEs (for example, ≥2% of subjects in any treatment group) will be summarized by preferred term. Serious TEAEs, TEAEs leading to discontinuation of IP and serious TEAEs leading to death, will be summarized by SOC, preferred term and treatment group.

For the summaries in parts A and B combined, patient-year exposure adjusted adverse event tables will be provided. Patient-year exposure will be calculated as (last day on treatment minus first day on treatment)/365.25.

Data for Part A includes all data with an onset date between date of first dose of IP up to last dose of Part A IP, inclusive, or 5 days, inclusive of last dose of IP for subjects that do not participate in Part B. Data for Part B includes all data with an onset date from first dose of Part B IP up to the last dose of Part B IP plus 5 days, inclusive. This rule is also applicable to Section 7.3, Section 7.4 and Section 7.5.

7.2 Adverse Events of Special Interest

Suicidal ideation (SI) and suicidal behavior (SIB) are the Adverse Events of Special Interest for this study.

7.3 Clinical Laboratory Variables

Laboratory data will be subjected to both a quantitative analysis (descriptive summary statistics) and qualitative analysis where frequencies of normal, above normal and below normal values will be computed.

Assessments will be included only when the subject is on treatment. All assessments will be used for listings. Data for Part A includes all data with an assessment date between date of first dose of Part A IP through last dose of last dose of Part A IP. Data for Part B includes all data with an assessment date from the first dose of Part B IP up to the last dose of Part B IP.

Hematology: Hemoglobin, hematocrit, red blood cell (RBC) count, white blood cell (WBC) count, differential WBC count, and platelet count

Biochemistry: Albumin, total protein, creatinine, creatine kinase, alkaline phosphatase (ALP), Aspartate Transaminase (AST), alanine aminotransferase, gamma glutamyltransferase, total bilirubin, direct bilirubin, lactate dehydrogenase, total cholesterol, triglycerides, glucose, chloride, calcium, potassium, prolactin, sodium, magnesium, bicarbonate, phosphorus, urea, uric acid

Urinalysis: Protein, glucose, blood and pH

Clinical laboratory test values are considered potentially clinically significant (PCS) if they meet either the low or high limits (PCS criteria) as listed in Table 1.

Table 1. Criteria for Potentially Clinically Significant Laboratory Tests

Parameter	SI Unit	Sex	Age Range	Lower Limit	Higher Limit
Biochemistry					
Albumin	g/L	F	5-17 years >17 years	<33 <30	- -
		M	5-14 years >14 years	<33 <30	- -
ALP	IU/L			-	>=1.5 x ULN
ALT	U/L			-	>= 3 x ULN
AST	U/L			-	>= 3 x ULN
Bicarbonate	mmol/L			<10	>40
Bilirubin Total	µmol/L			-	>= 2 x ULN
Blood Urea Nitrogen	mmol/L			-	>35.71
Calcium	mmol/L			<1.7465	>3.11875
Chloride	mmol/L			<80	>120
Cholesterol	Mmol/L			-	>7.77
Creatine kinase	U/L			-	>1.5 x ULN
Creatinine	µmol/L	F M 	31-364 days 1-3 years 4-6 years 7-9 years 10-13 years 14-15 years >=16 years <= 3 years 4-6 years 7-9 years 10-11 years 12-13 years 14-15 years >=16 years	- - - - - - - - - - - - - -	>159.156 >185.682 >212.208 >185.682 >212.208 >238.734 >251.997 >185.682 >212.208 >159.156 >185.682 >212.208 >291.786 >310.3542
eGFR/1.73m2 (Schwartz)	mL/min			<30	-
GGT	U/L			-	>1.5 x ULN
Glucose	mmol/L		8-364 days >=1 year	<2.4975 <2.22	>13.875 >16.65
LDH	U/L			-	> 2 x ULN
Magnesium	mmol/L			<0.41152	> 1.2
Potassium	mmol/L			<3.0	> 6

Table 1. Criteria for Potentially Clinically Significant Laboratory Tests

Parameter	SI Unit	Sex	Age Range	Lower Limit	Higher Limit
Phosphate	mmol/L			<0.6458	>2.2603
Sodium	mmol/L			<130	>155
Total protein	g/L			<54	>86
Triglycerides	U/L			-	>2.5 x ULN
Uric acid	µmol/L	F		-	>532
		M		-	>591
Hematology					
Hematocrit	V/V			<0.2	>0.6
Hemoglobin	g/L			<80	>200
Platelet	x 10^9/L			<50	>600
White Blood Cell Count	x 10^9/L			<2	>35
Red blood cells	x 10^{12}/L			<3.0	-
Neutrophils/leukocytes	%		2-5 years 6-11 years 12-17 years >17 years	<20 <25 <30 <40	>65 >65 >70 >75
Lymphocytes/leukocytes	%		2-5 years 6-11 years 12-17 years >17 years	<12 <12 <12 <15	>63 >57 >52 >47
Monocytes/leukocytes	%			-	>15
Eosinophils/leukocytes	%			-	>10
Basophils/leukocytes	%			-	>10

ALT: alanine aminotransferase; ALP: alkaline phosphatase; AST: aspartate aminotransferase; LLN: Lower limit of normal value provided by the laboratory; ULN: Upper limit of normal value provided by the laboratory

The following analyses will be performed:

- Standard descriptive summary statistics at each scheduled measuring time point and the last individual measuring time point.
- Standard descriptive summary statistics for the absolute change from baseline to each scheduled measuring time point after baseline and the last individual measuring time point.
- The number and percentage of subjects with post-baseline PCS values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with available baseline values and at least 1 post-baseline assessment. The numerator is the total number of subjects with at least 1 post-baseline PCS value.
- Shift tables displaying shifts from baseline with respect to the PCS limits between baseline and each scheduled measuring time point after baseline and the last individual measuring time point will be provided.

A supportive listing of subjects with post-baseline PCS values will be provided including the subject number, site, baseline, and post-baseline values.

7.4 Vital Signs

Descriptive statistics for vital signs (eg., systolic and diastolic blood pressure, pulse rate, and body weight) and their changes from baseline at each post-baseline visit and at the end of study will be presented by treatment group.

Vital sign values will be considered PCS if they meet both the observed value criteria and the change from baseline criteria listed in Table 2.

Table 2. Criteria for Potentially Clinically Significant Vital Signs

Vital Sign Parameter	Subject Age	Flag	Criteria[a]	
			Observed Value	Change from Baseline
Systolic blood pressure (mmHg)		High	>110 (0 to 5 years) >120 (>5 to 10 years) >130 (>=11 years)	Increase of ≥20
		Low	≤90	Decrease of ≥20
Diastolic blood pressure (mmHg)		High	>75 (0 to 5 years) >80 (>5 to 10 years) >90 (>=11 years)	Increase of ≥15
		Low	≤50	Decrease of ≥15
Pulse rate (beats per minute)	<1 year 1-2 years 3-6 years >=7 years <=4 years >=5 years	High	>190 >160 >150 >140 <50 <40	Increase of ≥15 Decrease of ≥15
Weight (kg)		High	-	Increase of ≥7%
		Low	-	Decrease of ≥7%

[a] A post-baseline value is considered as a PCS value if its meets either criteria for observed value or change from baseline.

The number and percentage of subjects with PCS post-baseline values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with baseline and at least 1 post-baseline assessment. The numerator is the total number of subjects with at least 1 PCS post-baseline vital sign value. A supportive listing of subjects with post-baseline PCS values will be provided including the subject number, site, baseline, and post-baseline PCS values.

The SBP and DBP normal ranges for children are provided in Table 3. The variable 'AGE' from the DM domain can be used in defining the normal ranges.

Table 3. Upper and Lower Limits for Blood Pressure Levels by Age and Sex

Age (year)	Boys				Girls			
	SBP (mmHg)		DBP (mmHg)		SBP (mmHg)		DBP (mmHg)	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
≤1	87	106	53	58	87	107	53	60
2	95	110	53	63	95	109	53	65
3	95	113	53	67	95	110	53	69
4	95	115	53	71	95	112	53	72
5	95	116	53	74	95	113	53	74
6	95	117	53	76	95	115	53	76
7	97	119	57	78	97	116	57	77
8	97	120	57	80	97	118	57	78
9	97	121	57	81	97	120	57	79
10	97	123	57	82	97	122	57	80
11	97	125	57	82	97	124	57	81
12	97	127	57	83	97	126	57	82
13	97	130	57	83	97	128	57	83
14	97	132	57	84	97	129	57	84
15	112	135	66	85	112	131	66	85
16	112	137	66	87	112	132	66	86
17	112	140	66	89	112	132	66	86

7.5 Electrocardiogram (ECG)

Descriptive statistics for ECG variables (eg, heart rate, PR interval, QRS interval, QT interval, and QTc interval) and their changes from baseline at each assessment time point will be presented by treatment group. QTc interval will be calculated using both Bazett (QTcB=QT/(RR)1/2) and Fridericia (QTcF=QT/(RR)1/3) corrections; and if RR is not available, it will be replaced with 60/hr in the correction formula. ECG interpretation will be summarized by visit. A shift table from baseline to each visit for qualitative ECG results will be presented.

Electrocardiogram variable values will be considered PCS if they meet or exceed the upper limit values listed in Table 4. The number and percentage of subjects with post-baseline PCS values will be tabulated by treatment group. The percentages will be calculated relative to the number of subjects with available non-PCS baseline and at least 1 post-baseline assessment. The numerator is the total number of subjects with at least 1 PCS post-baseline ECG value. A listing of all subjects with post-baseline PCS value will be provided including the subject number, site, baseline, and post-baseline PCS values.

Table 4. Criteria for Potentially Clinically Significant ECG Values

ECG Parameter	Subject Age	Unit	Higher Limit
QRS Interval	<4 years	msec	>90
QRS Interval	4-17 years	msec	>100
QRS Interval	<4-17 years	msec change from baseline	>=60
PR Interval	<4-17 years	msec	200
QTcF Interval	<4-17 years	msec	450
QTcF Interval	<4-17 years	msec change from baseline	>=60

8. PHARMACOKINETIC ANALYSIS

For subjects who consent to the sparse PK sampling, individual plasma concentrations with corresponding sampling times post-dose will be tabulated per visit. A graphical presentation of the average concentration over time by dose will be used to visualize the plasma concentration-time relationship. Descriptive statistics will be calculated per visit by dose at specific post-dose time intervals. Plasma levels below the limit of quantification will be flagged.

Plasma concentrations measured in this study will be combined with the PK data from previous trials in pediatric subjects and a population PK analysis will be performed on the pooled data. The relationship between PK and efficacy/safety endpoints of interest may be explored. This analysis will be subject of a separate report.

9. PHARMACODYNAMICS ANALYSIS

Not applicable

10. OTHER ANALYSES

During the Placebo-controlled Part [Part A] of the study, global evaluations and QoL Health economics and outcomes research endpoints will be assessed. These include:

- PGI-S.
- CGI-S.
- PedsQL™ Gas and Bloating Module for subjects with functional gastrointestinal disorders.

The PGI-S and CGI-S will jointly be used in an anchor-based analysis to determine a clinically meaningful within-patient change threshold for change in the number of SBMs. Baseline will be defined as the PGI-S/CGI-S from screening (ie, prior to disimpaction). Subjects with assessments at both baseline and Week 12 will be classified, regardless of treatment assignment, by the magnitude of the change categories between baseline and Week 12; specifically, the categories will be (- indicating improvement and + indicating deterioration): -3 categorical shift, -2 categorical shift, -1 categorical shift, 0 categorical shift, +1 categorical shift, +2 categorical shift, +3 categorical shift. Shift tables from baseline to Week 12 will be presented to show the changes for PGI-S and CGI-S, respectively. In the event categories on the extreme (eg, -3 or +3) have zero or <10 subjects then those categories will be collapsed or pooled with the neighboring categories.

Once classified the following anchor-based empirical cumulative distribution function (eCDF) and probability density function (PDF; many times estimated using kernel density estimation) curves, including the sample size and median score for each eCDF and PDF anchor curve for the primary endpoint; along with the sample size, and the 10th, the 25th, the 50th (median), the 75th, and the 90th percentile of the primary endpoint. In the event categories are collapsed then supplemental eCDF and PDF will be provided using the un collapsed categories.

Baseline for the PedsQL Gas and Bloating Module will be the screening assessment prior to disimpaction. For both baseline and postbaseline the total scale score is calculated as the sum of all items divided by the total number of nonmissing items. The subscale score is calculated as the sum of subscale items divided by the total number of non-missing subscale items. If more than 50% of the items in the (sub)scale are missing, the score for this (sub)scale will be set to "missing".

PedsQL Gas and Bloating Module will be summarized as follows:

- Absolute values and change from baseline.
- Standardized Response Mean (SRM).

- The SRM is defined as the mean change from baseline divided by the SD of the change. Assessments performed
- The SRM will also be presented classified: <0.2, 0.2-<0.5, 0.5-<0.8, and ≥0.8, with classes regarded as no effect, small, moderate, and large, respectively.
- Absolute values.
- All global evaluations and QoL Health economics and outcomes research endpoints will be analyzed by each scheduled visit (ie, baseline, week 8, week 24). In addition, the final on-treatment assessment, defined as either the final scheduled time point for completers or the early discontinuation visit for early terminators, will be summarized. Descriptive statistics for each subscale, will be tabulated per treatment group over the mITT analysis set. For continuous endpoints, each of the two treatment arms will be compared with placebo using the analysis of covariance with the age and baseline number of SBMs/week as covariates. For categorical endpoint (only classification of the SRM), treatment arms will be compared using the CMH test adjusted for age and baseline number of SBMs/week.

11. INTERIM ANALYSIS/DATA MONITORING COMMITTEE

Two IA will be conducted during the study.

The primary objective of the first IA is to evaluate stopping the study for futility. The first IA will be conducted when 50% of target number of toilet-trained subjects who are at least 3 years of age (ie, 120 subjects) are randomized into the study and have either completed or had withdrawn from the Placebo-controlled part (Part A). Futility evaluation will be based on the conditional power (Lan and Wittes 1988) using a stopping threshold of 20% for each dose arm.

Only if both comparisons, low dose and high dose versus placebo, based on the primary endpoint have a conditional power (probability) of less than 20% the study will be stopped for futility. In case of study termination due to futility, no further inferential efficacy analysis will be conducted since the original power calculation was based on 240 subjects (80 per each arm of placebo, low dose and high dose) who should have either completed or withdrawn from part A. Inferential analyses on partial data of patients at varying time spent in part A is not meaningful anymore. Instead, all efficacy endpoints including primary, secondary and exploratory endpoints stated in Section 2 and Section 6 will be reported in a descriptive manner. Health economics and outcomes research endpoints stated in Section 10 will be reported descriptively excluding analyses for anchor-based distributions.

In all other situations the study will continue as planned, with both high and low dosing arms and the placebo arms.

Without loss of generality, the conditional power for low dose is calculated based on B-value (Lan and Wittes 1988). The formula for conditional power is expressed as:

$$C_p(t) = 1 - \Phi\left\{\frac{Z_\alpha - B(t)/t}{\sqrt{1-t}}\right\}$$

The B-value is shown as:

$$B(t) = Z_n\sqrt{t}$$

$t = n/N$ represent the proportion of the information observed at the time of IA.

n is the sample size observed at the time of IA, N is the planned total sample size.

Zn is the test statistic observed at the time of IA.

$$Z_n = \frac{\bar{X}_k - \bar{Y}_m}{S_n\sqrt{k^{-1} + m^{-1}}}$$

$\bar{X}_k$ the means of response after imputation and k is the sample size for low dose treatment group at time of IA.

$\bar{Y}_m$ the means of response after imputation and m is the sample size for placebo group at time of IA,

Sn is the pooled estimate of σ.

To make the first IA be representative of the primary analysis, twenty imputations will be performed. The average of those twenty Z_n obtained from each imputed dataset is used for calculating $B(t)$ and then the conditional power $C_p(t)$.

In the case where one treatment arm has less than 19 subjects with week 12 data, each day with missingness in the weeks where the subject has less than 4 diary days will be imputed as 0 so that we have 7 days for those weeks after imputation. In such case, we assume that subjects who are not filling out the edairy data daily are often not filling out the diary due to not having a bowel movement, therefore, imputing missing days as 0 will provide the most conservative approach. After being imputed with 0, the calculation of average number of SBM would follow the same steps shown in Section 6.1.

A second IA will be performed when all randomized and toilet-trained subjects have completed or have withdrawn from the Placebo-controlled Part A. At this IA, the primary endpoint will be analyzed as planned. The DMC will then make a recommendation to the sponsor to stop or continue the remainder of the study. If one of the doses shows a statistically significant difference (using the Hochberg step-up procedure for adjustment of multiplicity) with placebo in primary endpoints (H1 or H2), the study will be continued.

To maintain data integrity, the results from the IAs based on unblinded data will be accessible only to the independent and external DMC, such that they can make a recommendation to the sponsor. At all points, the study team directly involved in the conduct of the study will remain blinded to maintain the double-blind nature of the study. All outputs for the interim analysis will be prepared by an independent analysis team. More details will be explained in the charter for the DMC.

More details will be explained in the charter for the DMC.

11.1 Data Monitoring Committee

An independent DMC will be installed to monitor accumulating safety data in this study including SI/SIBs to detect evidence of possible safety issues. (Additional information on the referral of SI/SIB to the DMC is provided in Section 7.2). The DMC will also be involved in both planned interim analyses to assess if efficacy is sufficient to continue the study. The DMC is independent and will give recommendation to the sponsor with regards to the termination of study. Ideally, the DMC will include 3 clinicians (of which at least one is a pediatrician), a PK expert, and a statistician. The exact composition will be described in the DMC charter.

Unblinded data will be at the disposal of the DMC. The DMC review will consist of a closed session and an open session. During the closed session, the DMC will look at the data (blinded and if needed unblinded) and come to a recommendation for the sponsor. During the open session, the DMC recommendation is discussed with the sponsor (blinded). The sponsor remains blinded. The composition of the DMC, the responsibilities of all DMC members, and the relation between DMC and sponsor will be described in more detail in a separate DMC charter. The DMC chair or sponsor can request additional analyses or meetings. The DMC chair can invite additional internal or external experts to discuss specific issues.

For safety monitoring, the DMC will meet every 6 months. A DMC meeting can be cancelled when, compared to the previous DMC meeting, less than 15 additional subjects are randomized.

12. DATA HANDLING CONVENTION

12.1 Definition of Baseline

Baseline for all efficacy analyses is defined as the last observed value for the efficacy assessment prior to taking the first dose of IP (based on dates or date/times).

12.2 Analysis Time Points

12.2.1 Relative Number of Days

The relative day is calculated as follows:

- Visit date – reference date + 1 day, when the visit date is on or after the reference date.
- Visit date – reference date, when the visit date is before the reference date.

The reference date of Study Part A is the day of first double-blind IP intake which by definition has DY=1. The reference date for Study Part B is the day of first IP intake post re-randomization. There is no DY=0.

12.2.2 Algorithm of Allocating Visits to Time Windows

All visits in the double-blind treatment period (including unscheduled visits but excluding visits without data) will be placed into time windows according to their relative day in the study, according to Table 5. Analysis windows for e-dairy data for the primary efficacy endpoint is shown in Table A3. Weekly fecal incontinence analysis windows will be defined as the assessment window in the Schedule of Activities in Appendix 1.

Table 5. Algorithm of Allocating Visits to Double-blind Time Windows

Time point label	Target day	Interval lower bound	Interval upper bound
Screening	1	-∞	1
Week 2	15	2	21
Week 4	29	22	43
Week 8	57	44	71
Week 12	85	72	99
Week 16	113	100	141
Week 24	169	142	197
Week 32	225	198	253
Week 40	281	254	309
Week 48	337	310	365
>Week 48	-	366	+∞
Final on treatment assessment	Last non-missing post-baseline double-blind assessment.		

If more than 1 visit is allocated into the same time window, only the visit with a relative day (DY) closest to the target day will be selected for analysis tables and figures. The non-selected visit(s) will only be listed. If there are multiple visits at the same distance of the scheduled visit day (meaning: equal ABS [DY target day]), then the 1 latest in time is selected. In case of screening, the last pre-baseline measurement is selected for analysis.

12.2.3 Analysis Periods

Analysis period	Start period	End period
Run-in	Date of signing the ICF	First treatment administration date in double-blind period -1 day
Double-blind (Week 1-12) *	First treatment administration date in double-blind period	Last visit in DB treatment phase (Part A)

Note that the last analysis period in case of early termination will always be ended by the study termination date (date of last contact) +5 days (5 x half-life of 24 hours).

12.3 General Data Reporting Conventions

Unless otherwise specified, continuous variables will be summarized using the following descriptive statistics: n, mean, median, SD, minimum, maximum. Categorical and count variables will be summarized by the number of subjects (n) and the percent of subjects in each category. Percentages will be presented with 1 decimal.

12.4 Repeated or Unscheduled Assessments of Safety Parameters

If a subject has repeated assessments before the start of IP, then the results from the final assessment made prior to the start of IP will be used as baseline. If end of study assessments is repeated or unscheduled, the last post-baseline assessment will be used as the end of study assessment for generating descriptive statistics. However, all post-baseline assessments will be used for PCS value determination and all assessments will be presented in the data listings.

12.5 Missing Date of IP

When the date of the last dose of IP is missing for a subject in the Safety Analysis Set, all efforts should be made to obtain the date from the investigator. See Section 12.6.5 for more information.

12.6 Missing Date Information for Prior or Concomitant Medications

For prior or concomitant medications, including rescue medications, incomplete (ie, partially missing) start date and/or stop date will be imputed. When the start date and the stop date are both incomplete for a subject, impute the start date first.

12.6.1 Incomplete Start Date

The following rules will be applied to impute the missing numerical fields. If the stop date is complete and the imputed start date is after the stop date, then the start date will be imputed using the stop date.

12.6.2 Missing Day and Month

- If the year of the incomplete start date is the same as the year of the date of the first dose of IP, then the day and month of the date of the first dose of IP will be assigned to the missing fields.
- If the year of the incomplete start date is before the year of the date of the first dose of IP, then December 31 will be assigned to the missing fields.
- If the year of the incomplete start date is after the year of the date of the first dose of IP, then 01 January will be assigned to the missing fields.

12.6.3 Missing Month Only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

12.6.4 Missing Day Only

- If the month and year of the incomplete start date are the same as the month and year of the date of the first dose of IP, then the day of the date of the first dose of IP will be assigned to the missing day.
- If either the year is before the year of the date of the first dose of IP or if both years are the same, but the month is before the month of the date of the first dose of IP, then the last day of the month will be assigned to the missing day.
- If either the year is after the year of the date of the first dose of IP or if both years are the same, but the month is after the month of the date of the first dose of IP, then the first day of the month will be assigned to the missing day.

12.6.5 Incomplete Stop Date

The following rules will be applied to impute the missing numerical fields. If the date of the last dose of IP is missing, then replace it with the last visit date. If the imputed stop date is before the start date (imputed or non-imputed start date), then the imputed stop date will be equal to the start date.

12.6.6 Missing Day and Month

- If the year of the incomplete stop date is the same as the year as of the date of the last dose of IP, then the day and month of the date of the last dose of IP will be assigned to the missing fields.
- If the year of the incomplete stop date is before the year of the date of the last dose of IP, then 31 December will be assigned to the missing fields.
- If the year of the incomplete stop date is after the year of the date of the last dose of IP, then 01 January will be assigned to the missing fields.

12.6.7 Missing Month Only

- The day will be treated as missing and both month and day will be replaced according to the above procedure.

12.6.8 Missing Day Only

- If the month and year of the incomplete stop date are the same as the month and year of the date of the last dose of IP, then the day of the date of the last dose of IP will be assigned to the missing day.
- If either the year is before the year of the date of the last dose of IP or if both years are the same, but the month is before the month of the date of the last dose of IP, then the last day of the month will be assigned to the missing day.
- If either the year is after the year of the last dose of IP or if both years are the same, but the month is after the month of the date of the last dose of IP, then the first day of the month will be assigned to the missing day.

12.7 Missing Date Information for Adverse Events

For AEs, the default is to only impute incomplete (ie, partially missing) start dates. Incomplete stop dates may also be imputed when calculation of the duration of an AE is required per the protocol. If imputation of an incomplete stop date is required, and both the start date and the stop date are incomplete for a subject, impute the start date first.

12.7.1 Incomplete Start Date

Follow same rules as in Section 12.6.1.

12.7.2 Incomplete Stop Date

When required per the protocol, follow the same rules as in Section 12.6.5.

12.8 Missing Severity Assessment for Adverse Events

If severity is missing for an AE starting prior to the date of the first dose of IP, then a severity of "Mild" will be assigned. If the severity is missing for an AE starting on or after the date of the first dose of IP, then a severity of "Severe" will be assigned. The imputed values for severity assessment will be used for incidence summaries, while the actual values will be used in data listings.

12.9 Missing Relationship to Investigation Product for Adverse Events

If the relationship to IP is missing for an AE starting on or after the date of the first dose of IP, a causality of "Related" will be assigned. The imputed values for relationship to double-blind IP will be used for incidence summaries, while the actual values will be presented in data listings.

13. ANALYSIS SOFTWARE

Statistical analyses will be performed using Version 9.3 (or newer) of SAS in a suitably qualified environment.

14. CHANGES TO ANALYSIS SPECIFIED IN PROTOCOL

14.1 Changes from Protocol

Not applicable.

14.2 Changes from Version 1 of SAP

Changes from version 1 to version 2 of the SAP are provided here, removed text is indicated using strikeout font, additions with bold fond. Editorial type changes, ie, spelling, grammar and formatting are not included.

Table 6. Changes Made in Version 2 of SAP

Section	Description of Change	Reason for Change
4.1.6	~~Incorrect timing of assessments: A primary/key secondary endpoint assessment performed outside a ±3 day window for the nominal visit date will be classified as a major protocol deviation.~~	This criteria will not apply, as for these endpoints, data is collected daily, and endpoint are averages over each week.
5.7	Investigational product dosing compliance for a specified period is defined as the **total number of days that the subject or caregiver report in the e-diary that the medication was taken by a subject during that period divided by the expected number of days medication to be taken in the same time period** ~~total number of tablets or liquid volume actually taken during the same period~~ multiplied by 100. ~~The total number of tablets or liquid volume actually taken is calculated by the total number of tablets or liquid volume dispensed minus the number of tablets or liquid volume returned.~~ Descriptive statistics for IP compliance will be presented by treatment group for each period between 2 consecutive visits as well as for the whole double-blind evaluation phase for the study for the Safety Analysis Set. These will include summary statistics (n, mean median, std, min and max) as well as categorization in appropriate categories (eg, <80%, 80-120%, > 120%). The above summary will be produced for each Study Part and toilet-trained status.	The calculation was changed to reflect the data collected.
6.1	If the UN covariance structure fails to converge, as an alternative, the Compound symmetry (CS) covariance matrix will be applied. ~~Degrees of freedom will be estimated using the Kenward-Roger approximation. For purpose of convergence, pooling of smaller centers might be needed, decisions on pooling will be made before unblinding.~~	
6.1.1	MMRM analysis with additional covariate and factors: The proposed MMRM model for primary efficacy analysis, including additional factors (center, tablet vs oral solution, global severity score) and covariate age. Note: When age is analyzed as a covariate, age group will not be included as a factor in the MMRM; **for purpose of convergence, pooling of smaller centers might be needed. Centers with fewer than 5 subjects in the Full Analysis Set will be**	

Table 6. Changes Made in Version 2 of SAP

Section	Description of Change	Reason for Change
	pooled into 1 of 4 regional pools. The pooling strategy will be reassessed and documented prior to unblinding to ensure that the regional pools are not over representative, ie, contain more subjects than the large enrolling centers, wherein large regional pools will be divided with centers being selected at random.	
7	For toilet-trained subjects, all safety data will be summarized for the whole study (Part A and Part B combined), and also separately for each Part A and Part B, unless otherwise specified. Part A will be summarized using the placebo, low dose, high dose and Motegrity overall. Part B will be summarized using the groupings of low dose, high dose and overall. Summaries of Part A and Part B are combined using the groupings placebo, placebo-low dose, placebo-high dose, low dose-low dose, high dose-high dose and Motegrity overall, unless otherwise specified. **Subjects randomized to Placebo in Part A who discontinued from the study prior to starting Part B will also be included in placebo-low dose or placebo-high dose.**	Clarification of planned analyses.
7.1	**Data for Part A includes all data with an onset date between date of first dose of IP up to last dose of Part A IP, inclusive, or 5 days, inclusive of last dose of IP for subjects that do not participate in Part B. Data for Part B includes all data with an onset date from first dose of Part B IP up to the last dose of Part B IP plus 5 days, inclusive. If subject is missing first dose of Part A from EC_ORAL form, it will be assumed that Part A first dose is baseline visit date. If subject has reached visit 14 and is missing last dose date of Part A on EC_ORAL form then Part A last dose will be assumed to be visit 14 date. If subject has reached visit 14 and is missing first dose of Part B it will be assumed that first dose Part B date is Visit 14 date + 1. If subject is missing Part B last dose date from EC_ORAL form and study day is greater than visit 14 date then it will be assumed that subject is still on Part B drug. The rules in this paragraph also apply to vital signs and ECG.**	Clarify the reporting of AEs between study Parts
0	**Assessments will be included only when subject is on treatment. All assessments will be used for listings. Data for Part A includes all data with an assessment date between date of first dose of Part A IP through last dose of last dose of Part A IP. Data for Part B includes all data with an assessment date from first dose of Part B IP up to the last dose of Part B IP.**	
12.2.2	Corrections to table 5 entries.	Entries in table 5 were corrected to reflect that first day of IP is counted as study day 0, previously it was counted as day 1.

Table 6. Changes Made in Version 2 of SAP

Section	Description of Change	Reason for Change
4.1	Combined Toilet-trained and Non-toilet-trained populations	Non-toilet-trained subjects will be pooled with toilet-trained for demographic and safety summary tables. It has been explicitly noted throughout where they will be shown separately.

14.3 Changes in Version 4 of SAP

Changes from version 3 to version 4 of the SAP are provided here, removed text is indicated using strikeout font, additions with bold font. Editorial type changes, ie, spelling, grammar and formatting are not included.

Table 7. Changes Made in Version 4 of SAP

Section	Description of Change	Reason for Change
2.2	The primary endpoint is the average change from baseline in number of spontaneous bowel movements (SBMs) per week derived from the (diary) data over 12 weeks, collected during the placebo-controlled part (Part A). Any bowel movement (BM) that occurs within 24 hours after intake of the rescue medication **or within 24 hours after disimpaction period** will not be considered as spontaneous.	Clarify for the in the case of baseline
4.1.6	Incorrect timing of assessments: A primary/key secondary endpoint assessment performed outside a ±3 day window for the nominal visit time will be classified as a major protocol deviation. **Note that this rule does not apply to ediary data, which are collected daily, and endpoint are averages over each week.**	Clarification
4.1.7	The Completers Analysis Set is applicable to Study Part A and will consist of all toilet-trained subjects who are at least 3 years of age in the mITT Analysis Set who have ~~an average number of~~ **an average number of SBMs available for all 12 weeks in the Placebo-controlled Part (Part A)** ~~completed the daily diary for at least 12 weeks during Part A~~ of the study. **The subject in this set must have at least 4 days with completed diary data in each of the 12 weeks** ~~to have availability~~ **to calculate the of average number of SBMs**.	Clarification
5.6	**If two sources of the end date are available the date collected on site will be used to determine last dose of IP**. ~~The last dose of IP will be determined for both Part A and Part B using drug accountability eCRF page along with the records within eDiary~~. **See Section 12.5 for missing last dose date**.	Change to reflect actual eCRF

Table 7. Changes Made in Version 4 of SAP

Section	Description of Change	Reason for Change
5.7	…**(difference in days as reported on the eCRF dosing form)**…	Clarification
6.1	…**or within 24 hours after disimpaction period**…	Clarification
7	**Subjects that did not reach Part B or did not participate in Part B and received low dose in part A will be included the low/low group. Subjects that did not reach Part B or did not participate in Part B and received high dose in part A will be included the high/high group. If a subject is in the safety set for Part A or Part B they will be included in the combined analyses.** Non toilet-trained subjects will be summarized similarly where data permit **together with toilet-trained subjects, unless specified otherwise.**	Clarification Non toilet-trained will be pooled with toilet-trained for majority of summary tables.
7.1	An AE (classified by preferred term) that occurs during this study will be considered a TEAE if it has a start date on or after the first dose of double-blind IP or if it has a start date before the date of the first dose of double-blind IP but increases in severity on or after the date of the first dose of double-blind IP. An AE that occurs more than 5 days (5 x half-life of 24 hours) after the date of the last dose of double-blind IP will not be counted as a TEAE. **If two sources of the end date of treatment are available, the date collected on site will be used to determine last dose of IP.**	Clarification
7.1	Data for Part A includes all data with an onset date between date of first dose of IP up to last dose of Part A IP, inclusive, or 5 days, inclusive of last dose of IP for subjects that do not participate in Part B. Data for Part B includes all data with an onset date from first dose of Part B IP up to the last dose of Part B IP plus 5 days, inclusive. **This rule is also applicable to Sections 7.3, 7.4 and 7.5**. ~~If subject is missing first dose of Part A from EC_ORAL form, it will be assumed that Part A first dose is baseline visit date. If subject has reached visit 14 and is missing last dose date of Part A on EC_ORAL form then Part A last dose will be assumed to be Visit 14 date. If subject has reached visit 14 and is missing first dose of Part B it will be assumed that first dose Part B date is Visit 14 date +1. If subject is missing Part B last dose date from EC_ORAL form and study day is greater than Visit 14 date then it will be assumed that subject is still on Part B drug. The rules in this paragraph also apply to vital signs and ECG.~~	Clarification
7.3	Assessments will be included only when subject is on treatment. All assessments will be used for listings. Data for Part A includes all data with an assessment date between date of first dose of Part A IP through last dose of last dose of Part A IP. Data for Part B includes all data with an assessment date from first dose of Part B IP up to the last dose of Part B IP. ~~If subject is missing first dose of Part A from EC_ORAL form, it will be assumed that Part A first dose is baseline visit date. If subject has reached visit 14 and is missing last dose date of Part A on EC_ORAL form then Part A last dose will be assumed to be Visit 14 date. If subject has reached visit 14 and is missing first~~	This has been replace.

Table 7. Changes Made in Version 4 of SAP

Section	Description of Change	Reason for Change
	~~dose of Part B it will be assumed that first dose Part B date is Visit 14 date +1. If subject is missing Part B last dose date from EC_ORAL form and study day is greater than Visit 14 date then it will be assumed that subject is still on Part B drug. The rules in this paragraph also apply to vital signs and ECG.~~	
7.4	Lower values added to Table 3	Lower values were missing
7.5	All safety analyses for non-toilet-trained subjects will be performed same as for toilet-trained subjects **and summarized together with those for toilet-trained subjects.** ~~Only selected TEAEs tables will be generated for non-toilet-trained subjects~~. The NTTSAF will be used for safety analysis	Reflect change to pool non toilet-trained with toilet trained.
10	– Anchor-based empirical cumulative distribution function (eCDF) with percentiles and probability density function (PDF) curves will be provided for each outcome endpoint, with the sample size included in each figure's legend. **Note that eCDF and PDF are only applicable for those endpoints (PGI-S, CGI-S and PedsQL) in this Section.**	Clarification
11	– $\bar{X}_k$ the ~~observed~~ means of response **after imputation** and k is the sample size for low dose treatment group at time of IA. – $\bar{Y}_m$ the ~~observed~~ means of response **after imputation** and m is the sample size for placebo group at time of IA, – Sn is the pooled estimate of σ. **To make the first IA be representative of the primary analysis, twenty imputations will be performed. The average of those twenty Z_n obtained from each imputed dataset is used for calculating $B(t)$ and then the conditional power $C_p(t)$.** **In the case where one treatment arm has less than 19 subjects with week 12 data, each day with missingness in the weeks where the subject has less than 4 diary days will be imputed as 0 so that we have 7 days for those weeks after imputation. In such case, we assume that subjects who are not filling out the edairy data daily are often not filling out the diary due to not having a bowel movement, therefore, imputing missing days as 0 will provide the most conservative approach.. After being imputed with 0, the calculation of average number of SBM would follow the same steps shown in Section 6.1.**	
12.2.2	Table 5	Days in table 5 were shifted to correct that first day of study is called 0 rather than 1.

Table 7. Changes Made in Version 4 of SAP

Section	Description of Change	Reason for Change
12.5	When the date of the last dose of IP is missing for a subject in the Safety Analysis Set, all efforts should be made to obtain the date from the investigator. **See section 12.6.5 for more information.** ~~If it is still missing after all efforts, then the latter of the last dispensing date or the last visit date when IP was returned will be used in the calculation of treatment duration.~~	

14.4 Changes in Version 5 of SAP

Changes from version 4 to version 5 of the SAP are provided here, removed text is indicated using strikeout font, additions with bold font. Editorial type changes, ie, spelling, grammar and formatting are not included.

Table 8. Changes Made in Version 5 Of SAP

Section	Description of Change	Reason for Change
2.3.2, 6.2	Update secondary endpoints to describe subgroups by reporter status as per protocol	Made consistent with Protocol
4.1	Combined Toilet-trained and Non-toilet-trained populations	Non-toilet-trained subjects will be pooled with toilet-trained for demographic and safety summary tables. It has been explicitly noted throughout where they will be shown separately.
5.2, 5.4, 5.5, 7.5	All safety analyses for non-toilet-trained subjects will be performed same as for toilet-trained subjects **and summarized together with those for toilet-trained subjects.**	Non-toilet-trained subjects will be pooled with toilet-trained for demographic and safety summary tables. It has been explicitly noted throughout where they will be shown separately.

Table 8. Changes Made in Version 5 Of SAP

Section	Description of Change	Reason for Change
11	**Only if both comparisons, low dose and high dose versus placebo, based on the primary endpoint have a conditional power (probability) of less than 20% the study will be stopped for futility. In case of study termination due to futility, no further inferential efficacy analysis will be conducted since the original power calculation was based on 240 subjects (80 per each arm of placebo, low dose and high dose) who should have either completed or withdrawn from part A. Inferential analyses on partial data of patients at varying time spent in part A is not meaningful anymore. Instead, all efficacy endpoints including primary, secondary and exploratory endpoints stated in Section 2 and Section 6 will be reported in a descriptive manner. Health economics and outcomes research endpoints stated in Section 10 will be reported descriptively excluding analyses for anchor-based distributions.**	Update to show analyses not being done due to futility.

15. REFERENCES

Dmitrienko, A., Kordzakhia, G. and Brechenmacher, T. 2016. Mixture-based gatekeeping procedures for multiplicity problems with multiple sequences of hypotheses. *J Biopharm Stat,* 26(4), 758-80.

Heitjan, D. F. and Little, R. J. A. 1991. Multiple imputation for the fatal accident reporting system. *Appl Statist,* 40(1), 13-29.

Lan, K. K. and Wittes, J. 1988. The B-value: a tool for monitoring data. *Biometrics,* 44(2), 579-85.

Rubin, D. B. 1987. *Multiple imputation for nonresponse in surveys.* New York: John Wiley & Sons, Inc.

Schenker, N. and Taylor, J. M. G. 1996. Partially parametric techniques for multiple imputation. *Comput Stat Data Anal,* 22(4), 425-46.

Appendix 1 Schedule of Activities

Table A1 Schedule of Activities – Placebo-controlled Part (Part A)

Period	Screening[a]	Placebo-controlled Part												
Visit	**1**	**2 Base-line**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Week	**Approx. -4**	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Study Day	**-33 to –1[b]**	**0**	**7**	**14**	**21**	**28**	**35**	**42**	**49**	**56**	**63**	**70**	**77**	**84**
Assessment window (in days)	NA	0	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3
On-site visit[c]	X	X				X				X				X
Telephone contact[d]	X[e]		X	X	X		X	X	X		X	X	X	
Informed consent/assent	X													
Eligibility assessment	X													
Randomization		X												
Re-randomization (placebo-subjects only)														X
Demography, medical and medication history[f]	X													
Assessment of Rome IV criteria[g]	X													X
Physical examination (complete)	X[h]													
Physical examination (targeted)		X[i]				X				X				X
Height	X	X				X				X				X
Weight	X	X				X				X				X
Vital signs	X	X				X				X				X
12-lead ECG	X[j]	X												X
Dispense investigational product		X				X				X				X[k]
Collect investigational product						X				X				X
Provide rescue medication[l,m]		X								X				
Provide disimpaction medication[m]	X													
Removal of fecal impaction[n]	X													
Dispense e-Diary	X													
Check e-Diary[o]	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Table A1 Schedule of Activities – Placebo-controlled Part (Part A)

Period	Screening[a]	Placebo-controlled Part												
Visit	**1**	**2 Base-line**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Week	**Approx. -4**	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Study Day	**-33 to –1[b]**	**0**	**7**	**14**	**21**	**28**	**35**	**42**	**49**	**56**	**63**	**70**	**77**	**84**
Assessment window (in days)	NA	0	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3
Confirm visit on eCOA device[p]	X	X				X				X				X
Behavioral therapy reminder	X	X				X				X				X
Concomitant medications check	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Concomitant surgeries/ procedures/psychiatric changes/CBT check	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Prohibited medications check	X	X	X	X	X	X	X	X	X	X	X	X	X	X
AE check[q]	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Fecal incontinence clinical assessment	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Retentive posturing clinical assessment[r]	X	X				X				X				X
Schedule date for next visit	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Laboratory Assessments														
Hematology	X					X								X
Serum chemistry	X					X								X
Urinalysis	X					X								X
Serum pregnancy test[s]	X													
Urine pregnancy test[s]		X				X				X				X
Sparse PK sampling (optional)[t]		X				X				X				X
Clinical Outcome Assessments														
e-Diary (including Wong Baker or NRS, BSFS)	X	◄———												———►
PGI-S and CGI-S	X	X				X				X				X
PedsQL™ GI (gas and bloating module)	X	X				X				X				X

Table A1 Schedule of Activities – Placebo-controlled Part (Part A)

Period	Screening[a]	Placebo-controlled Part												
Visit	**1**	**2 Base-line**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Week	**Approx. -4**	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Study Day	**-33 to –1[b]**	**0**	**7**	**14**	**21**	**28**	**35**	**42**	**49**	**56**	**63**	**70**	**77**	**84**
Assessment window (in days)	NA	0	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3

AE=adverse event; BSFS=Bristol Stool Form Scale; CBT=cognitive behavioral therapy; CGI-S=Caregiver Global Impression of Severity; ECG=electrocardiogram; eCOA=electronic clinical outcome assessment; eCRF=electronic case report form; e-Diary=electronic diary; ET=early termination; ICF=informed consent form; IP=investigational product; NRS=numerical response scale; PedsQL™ GI=Pediatric Quality of Life Inventory (PedsQL™) Gastrointestinal Symptoms; PEG=polyethylene glycol; PGI-S=Patient Global Impression of Severity; PK=pharmacokinetic(s)

[a] Since, in the screening period, the subject's diagnosis is checked and since all subjects need to be disimpacted during the screening period, the screening period includes at least 3 telephone contacts and can include an unscheduled visit.

[b] The screening period can be as short as 10 days, as long as 7 consecutive days of diary data have been recorded.

[c] All subjects should be instructed to withhold dosing on in-clinic visit days (Part A and Part B) until after their visit. This includes Baseline (Week 0), Week 4 (Visit 6), Week 8 (Visit 10), Week 12 (Visit 14), Week 16 (Visit 15b), Week 24 (Visit 17), Week 32 (Visit 19), Week 40 (Visit 21), and Week 48 (Visit 23/ET). Note that if a subject does take their daily dose of IP prior to the on-site visit, this will not result in a protocol deviation unless the subject is participating in PK sampling at Week 4 (Visit 6), Week 8 (Visit 10), and Week 12 (Visit 14). Subjects will be dosed on site on Baseline (Visit 0).

[d] Any telephone visit can be changed to an onsite visit at the investigator's discretion.

[e] Regular telephone contacts will be made with the subject/parent(s)/caregiver(s) to decide on the continuation of the screening period, the initiation of the disimpaction procedure, and the success or failure of the procedure.

[f] Information related to disease history will also be collected.

[g] Rome IV criteria will be assessed at the time of screening and at the Week 12 visit. If a subject discontinues from the study for any reason prior to completing Part A, Rome IV will be assessed at the time of the early termination visit.

[h] The screening visit (Visit 1) physical examination may include an optional rectal examination to confirm the presence or absence of fecal impaction. The anal and cremasteric reflex examination is also optional.

[i] The Visit 2 physical examination may also include an optional rectal examination to confirm the absence of fecal impaction.

[j] An ECG recording up to 6 months prior to the screening visit is also acceptable. The eligibility of the subject is based on the assessment of the ECG by the investigator at screening (Visit 1).

[k] Subjects will be rerandomized to Part B IP.

[l] Current laxatives will be discontinued and replaced by the protocol-specified rescue medication.

[m] Subjects will get a prescription for the rescue and/or disimpaction medication as well as instructions on how to obtain these medications. Rescue medications can be provided more or less frequently at the discretion of the investigator.

[n] The disimpaction can take 3 days if the PEG 3350 is effective the first time, and it can take as long as 12 days if a 3-6 day cycle of PEG needs to be repeated (see Protocol Section 6.7.3.2).

[o] The investigator or study staff will remind the subject and/or parent(s)/caregiver(s)/legally authorized representative(s) to complete the e-Diary every day, to transmit the data regularly, and to bring the device at the next visit.

[p] Visit confirmation can be done on eCOA portal for all visits other than screening.

Table A1 Schedule of Activities – Placebo-controlled Part (Part A)

Period	Screening[a]	Placebo-controlled Part												
Visit	**1**	**2 Base-line**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Week	**Approx. -4**	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Study Day	**-33 to –1[b]**	**0**	**7**	**14**	**21**	**28**	**35**	**42**	**49**	**56**	**63**	**70**	**77**	**84**
Assessment window (in days)	NA	0	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3	±3

[q] Adverse events will be reported from signing consent/assent onwards until the last study-related contact.
[r] The presence of any of these symptoms in the past 4 weeks will be recorded in the eCRF.
[s] A serum (screening)/urine (all other time points) pregnancy test will be performed for female subjects aged ≥12 years and/or subjects <12 years who have started menarche.
[t] This is an optional assessment and requires specific consent on the ICF. Sampling times between 1 to 3 hours post-dose after the first dose (on Day 0) and between 14 to 26 hours post-dose from the previous day's dosing at Week 4 (Visit 6), Week 8 (Visit 10), and Week 12 (Visit 14). At Week 4 (Visit 6), Week 8 (Visit 10), and Week 12 (Visit 14), dosing should be withheld until after the PK sample has been drawn.

Table A2 Schedule of Activities – Safety Extension Part (Part B)

Period	Safety Extension Part										Follow-up
Visit	**15a**	**15b**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23/ET**	**24**
Week	**14**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**
Study Day	**98**	**112**	**140**	**168**	**196**	**224**	**252**	**280**	**308**	**336**	**364**
Assessment Window (in days)	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**
Onsite Visit[a]		X		X		X		X		X	
Telephone Visit[b]	X		X		X		X		X		X
Targeted Physical examination		X		X		X		X		X	
Weight		X		X		X		X		X	
Vital signs		X		X		X		X		X	
12-lead ECG		X				X				X	
Dose increase based on weight				X [c]							
Dispense investigational product		X		X		X		X			
Collect investigational product		X		X		X		X		X	
Provide rescue medication [d]		X		X		X		X			
Check e-Diary[e]		X		X		X		X		X	
Return e-Diary										X	
Confirm visit on eCOA device[f]										X[g]	
Behavioral therapy reminder		X		X				X			
Check toilet-training status	X	X	X	X	X	X	X	X	X	X	
Concomitant medications check	X	X	X	X	X	X	X	X	X	X	X
Concomitant surgeries/procedures check/psychiatric changes/CBT check	X	X	X	X	X	X	X	X	X	X	X
Prohibited medications check	X	X	X	X	X	X	X	X	X	X	X
AE check[h]	X	X	X	X	X	X	X	X	X	X	X

Table A2 Schedule of Activities – Safety Extension Part (Part B)

Period	Safety Extension Part										Follow-up
Visit	**15a**	**15b**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23/ET**	**24**
Week	**14**	**16**	**20**	**24**	**28**	**32**	**36**	**40**	**44**	**48**	**52**
Study Day	**98**	**112**	**140**	**168**	**196**	**224**	**252**	**280**	**308**	**336**	**364**
Assessment Window (in days)	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**
Schedule/confirm date for next visit	X	X	X	X	X	X	X	X	X		
Clinical Laboratory Assessments											
Hematology		X				X				X	
Serum chemistry		X				X				X	
Urinalysis		X				X				X	
Urine pregnancy test[i]		X		X		X		X		X	
Clinical Outcomes Assessments											
e-Diary	←									→	

AE=adverse event; CBT=cognitive behavioral therapy; ECG=electrocardiogram; eCOA=electronic clinical outcome assessment; e-Diary=electronic diary; ET=early termination; IP=investigational product; PedsQL GI=Pediatric Quality of Life Inventory (PedsQL Gastrointestinal Symptoms)

[a] All subjects should be instructed to withhold dosing on in-clinic visit days (Part A and Part B) until after their visit. This includes Baseline (Week 0), Week 4 (Visit 6), Week 8 (Visit 10), Week 12 (Visit 14), Week 16 (Visit 15b), Week 24 (Visit 17), Week 32 (Visit 19), Week 40 (Visit 21), and Week 48 (Visit 23/ET). Note that if a subject does take their daily dose of IP prior to the on-site visit, this will not result in a protocol deviation unless the subject is participating in PK sampling at Week 4 (Visit 6), Week 8 (Visit 10), and Week 12 (Visit 14). Subjects will be dosed on site on Baseline (Visit 0).

[b] Any telephone visit can be changed to an onsite visit at the investigator's discretion.

[c] At Week 24 (Visit 17), each subject weighing <50 kg at baseline can undergo a dose adjustment for oral solution based on weight. In case the subject has crossed the 50-kg threshold, he/she will be switched from oral solution to tablet, provided he/she can swallow the tablet. If the subject cannot swallow the tablet, he/she can receive the tablet dose as oral solution. Depending on the treatment group, subjects cannot exceed the maximum dose of 2 or 4 mg.

[d] Subjects will get a prescription for the rescue medication as well as instructions on how to obtain these medications. Rescue medication can be provided more or less frequently at the discretion of the investigator.

[e] The investigator or study staff will remind the subject and/or parent(s)/caregiver(s)/legally authorized representative(s) to complete the e-Diary every day, to transmit the data regularly, and to bring the device at the next visit.

[f] Visit confirmation can be done on eCOA portal.

[g] For ET, not Visit 23.

[h] Adverse events will be reported from signing the TAK-555-3010 consent/assent onwards until the last study-related contact.

[i] A urine pregnancy test will be performed for female subjects aged ≥12 years and/or subjects <12 years who have started menarche.

Table A3 Analysis Window for E-dairy Data for Primary Efficacy Endpoint

Visit	Week	Study Day	Assessment window (in days)	
			lower limit	upper limit
2 - Baseline	0	0	-33	0
3	1	4	1	7
4	2	11	8	14
5	3	18	15	21
6	4	25	22	28
7	5	32	29	35
8	6	39	36	42
9	7	46	43	49
10	8	53	50	56
11	9	60	57	63
12	10	67	64	70
13	11	74	71	77
14	12	81	78	84

Note: Baseline values do not include the disimpaction period or 24 hours after the end of disimpaction period.

Appendix 2 Sample SAS Codes

```
**************************************************************************;
********** hybrid approach for imputation for primary efficacy**************;
**************************************************************************;
   proc sort DATA=sbm12weeks;
    by treat subjid age reason basesev week;
   run;

  proc transpose data =sbm12weeks out=xweeks(drop=_name_ _label_) prefix=wk;
   var chg;
    by treat subjid age reason basesev week;

    id week;
  run;

  proc mi data=xweeks out=impothers nimpute=20 seed=1010 noprint minimum = . . 0 0 0 0 0 0 0 0 0 0
0 0 0  k=5;
     where reason NOT IN ('ADVERSE EVENT','INEFFECTIVE');
      by treat ;
     VAR age basesev wk0-wk12;
      FCS REGPMM (wk0  = age basesev);
      FCS REGPMM (wk1  = wk0 );
      FCS REGPMM (wk2  = wk1 wk0 );
      FCS REGPMM (wk3  = wk2 wk1 wk0 );
      FCS REGPMM (wk4  = wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk5  = wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk6  = wk5 wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk7  = wk6 wk5 wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk8  = wk7 wk6 wk5 wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk9  = wk8 wk7 wk6 wk5 wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk10 = wk9 wk8 wk7 wk6 wk5 wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk11 = wk10 wk9 wk8 wk7 wk6 wk5 wk4 wk3 wk2 wk1 wk0 );
      FCS REGPMM (wk12 = wk11 wk10 wk9 wk8 wk7 wk6 wk5 wk4 wk3 wk2 wk1 wk0 );
    run;

**** use WOCF for subjects dropping out due to AE or lack of efficacy;

  data IMPAE;
   set xweeks;
   where reason in ('ADVERSE EVENT', 'INEFFECTIVE');
   array week[12] wk1-wk12;
    ** Set worst value to 0;
    WOCF=0
    ** Set Worst value to min for that subject.
```

```
     if week[1] ne . then WOCF= max(WOCF,week[1]);
     do i=2 to 12;
       if week[i] ne . then WOCF=min(week[1],week[i]);
     end;
    do i=1 to 12;
     if week[i] = .  then week[i]= WOCF;
    end;
    do _imputation_=1 to 20; output; end;
    drop i  WOCF;
   run;
**** combine;

   data xweeksimp;
    set impothers impae;
   run;

   proc sort data=xweeksimp;
     by _imputation_ treat subjid age basesev ;
   run;

**** transpose back so MMRM can be applied;;

   proc transpose data=xweeksimp out=weeksimp(rename=(col1=chg)) name=visit ;
    var wk1-wk12;
     by _imputation_ treat subjid age  basesev ;
   run;

   data weeksimp;
    set weeksimp;
    week=1*substr(visit,3,1);
   run;


**********************************************************************;
********** Calculate Primary Efficacy Endpoint After Hybrid Imputation*********;
**********************************************************************;

**** apply MMRM per imputation set;

   ods output diffs=diffs lsmeans=lsmeans ;
   PROC MIXED DATA=weeksimp noinfo noclprint;
    class treat subjid week ;
    model chg = base treat week treat*week age;
    repeated week /subject=subjid type=un;
    lsmeans treat/pdiff=control('_PLA') cl adjust=dunnett;
    by _imputation_;
```

```
run;

*** Combine results;
PROC MIANALYZE DATA=diffs;
 ODS OUTPUT PARAMETERESTIMATES=mi_mixed;
 MODELEFFECTS estimate;
 STDERR;
RUN;



**************************************************************************;
********** Calculating Conditional Power After Hybrid Imputation **********;
**************************************************************************;

Step 1: Calculate the mean and standard deviations by treatment groups within each
imputation dataset.
proc sort data=weeksimp;
by impnum trt USUBJID;
run;

/*calculate the average change in SBM across visit by Subject*/
proc means data=weeksimp;
by impnum trt USUBJID;
var chg;
output out=imp_means mean=meanchg stddev = sdchg;
run;

/*calculate the average change and standard deviation across Subject by treatment
group*/
proc means data=imp_means;
by impnum trt;
var meanchg;
output out=imp_means_all mean=meanchg stddev = sdchg;
run;

Step 2:Use the means, std deviations and sample size from each treatment group within
each imputation dataset to calculate the zn for each imputation dataset using the Lan
and Wittes (1988) method.
     Step 2a: Calculate the change in means(delta).
     Step 2b: Calculate the pooled standard deviation (psd)
     Step 3b: Calculate Zn.

/*compute zn for each imputation dataset*/
data zn_scores;
merge placebo PrucalopridеLowDose PrucalopridеHighDose;

/* compute overall sample size*/
n=n_placebo+n_low+n_high;

/*calculate zn for low dose versus placebo*/
deltalow=meanchg_low-meanchg_placebo; /*compute x-bar1 - x-bar2*/
psdlow=sqrt (( (n_low-1)*(sdchg_low**2)+ (n_placebo-
1)*(sdchg_placebo**2))/(n_low+n_placebo-2)); /*compute pooled standard deviation*/
```

```
zn_low=(meanchg_low-meanchg_placebo) / ( psdlow*(sqrt(n_low**-1 + n_placebo**-1)));
/*compute zn*/

/*calculate zn for high dose versus placebo*/
deltahigh=meanchg_high-meanchg_placebo;
psdhigh=sqrt (( (n_high-1)*(sdchg_high**2)+ (n_placebo-
1)*(sdchg_placebo**2))/(n_high+n_placebo-2));
zn_high=(meanchg_high-meanchg_placebo) / ( psdhigh*(sqrt(n_high**-1 + n_placebo**-
1)));
run;
```

Step 3: Calculate the average Zn across all imputations by group (high versus low).
***Note: could use proc mi to do this, however because we are only interested in the average and not in the variance we can also take an average.**

```
/*Calculate the average zn across all imputation datasets for each zn group(low and
high)*/

proc means data=zn_scores;
var zn_high zn_low n;
output out=average_zn ;
run;
```

Step 4: Calculate the conditional probability per Lan and Wittes (1988).

```
/*calculate conditional power*/
data cp_stats;
set average_zn ;
where _STAT_="MEAN";
t=n/240; /* Information time*/
bt_high=zn_high*sqrt(t); /*compute parameter B(t)/sqrt(t)*/
bt_low=zn_low*sqrt(t);
cp_high=1-probnorm((1.96 - (bt_high /t))/ (sqrt(1-t))) ; /*compute conditional power*/
cp_low=1-probnorm((1.96 - (bt_low /t))/ (sqrt(1-t))) ;
run;
```


Code For Sensitivity Analyses:

```
**********************************************************************;
*************** placebo multiple imputation**************************;
**********************************************************************;

proc mi data=xweeks out=impothers nimpute=40 seed=2019 noprint;
 where treat='_PLA';
 var wk1-wk8;
 mcmc outest=posteriors_pla; * Here we save the placebo posteriors;
run;

data posteriors_pla(type=est); * Here we assign the placebo posteriors;
   set posteriors_pla;
```

```
      if  0 < _imputation_ <= 20 then      treat='_PLA';
      if 20 < _imputation_ <= 40 then do;  treat='PRU ';
         _imputation_=_imputation_- 20;
      end;
   run;

   proc sort;
    by treat;
   run;

   proc mi data=xweeks out=im_placebo ;
      by treat;
      var wk1-wk8;
      mcmc inest=posteriors_pla; * Here we use the placebo posteriors;
   run;

   proc sort data=im_placebo;
     by _imputation_ base treat week age center sex formulation basesev race ;
   run;

**** transpose back so MMRM can be applied;;

   proc transpose data=im_placebo out=weeksimppla(rename=(col1=chg)) name=visit ;
    var wk1-wk8;
     by _imputation_ base treat week age center sex formulation basesev race;
   run;

   data weeksimppla;
    set weeksimppla;
    week=1*substr(visit,3,1);
   run;

**** apply MMRM per imputation set;

   ods output diffs=diffs_impla ;
   ods trace off;
   PROC MIXED DATA=weeksimppla noinfo noclprint;
    class treat subjid sex week race center formulation basesev ;
    model chg = base treat week age treat*week;
    repeated week /subject=subjid type=un;
    lsmeans treat/pdiff=control('_PLA') cl adjust=dunnett;
    by _imputation_;
   run;

   *** Combine results;
   PROC MIANALYZE DATA=diffs;
```

```
    ODS OUTPUT PARAMETERESTIMATES=mi_mixed;
    MODELEFFECTS estimate;
    STDERR;
   RUN;
*************************************************************************;
******************** on-treatment multple imputation ******************;
*************************************************************************;

   proc mi data=xweeks out=mi_ontreat nimpute=20 seed=2019 noprint;
    var wk1-wk8;
    by treat;
    mcmc;
   run;

   proc sort data=mi_ontreat;
     by _imputation_ base treat week age center sex formulation basesev race;
   run;

**** transpose back so MMRM can be applied;;

   proc transpose data=mi_ontreat out=weeksimp_ot(rename=(col1=chg)) name=visit ;
    var wk1-wk8;
     by _imputation_ base treat week age center sex formulation basesev race;
   run;

   data weeksimp_ot;
    set weeksimp_ot;
    week=1*substr(visit,3,1);
   run;

**** apply MMRM per imputation set;

   ods output diffs=diffs_ot ;
   ods trace off;
   PROC MIXED DATA=weeksimp_ot noinfo noclprint;
    class treat subjid sex week race center formulation basesev ;
    model chg = base treat week age treat*week;
    repeated week /subject=subjid type=un;
    lsmeans treat/pdiff=control('_PLA') cl adjust=dunnett;
    by _imputation_;
   run;

   *** Combine results;
   PROC MIANALYZE DATA=diffs_ot;
    ODS OUTPUT PARAMETERESTIMATES=mi_mixed;
```

```
 MODELEFFECTS estimate;
 STDERR;
RUN;
```