Efficacy, Safety, and Tolerability of Fosmetpantotenate (RE-024), a Phosphopantothenate Replacement Therapy, in Patients with Pantothenate Kinase-Associated Neurodegeneration (PKAN): A Randomized, Double-Blind, Placebo-Controlled Study With an Open-Label Extension

Unique Protocol ID:	024PKAN15004
NCT Number:	NCT03041116
EudraCT Number:	2016-001955-29
Date of SAP:	22 April 2019

Statistical Analysis Plan

EFFICACY, SAFETY, AND TOLERABILITY OF FOSMETPANTOTENATE (RE-024), A PHOSPHOPANTOTHENATE REPLACEMENT THERAPY, IN PATIENTS WITH PANTOTHENATE KINASE-ASSOCIATED NEURODEGENERATION (PKAN): A RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED STUDY WITH AN OPEN-LABEL EXTENSION

Protocol Number: 024PKAN15004

IND #: 121943
EudraCT #: 2016-001955-29

Version 2.0

Date: **22 April 2019**

SPONSOR SIGNATURE PAGE

Approved by: PPD

02 MAY 2019
Date

Retrophin, Inc.

PPD

02 MAY 2019
Date

Retrophin, Inc.

PPD

02 May 2019
Date

Retrophin, Inc.

PAREXEL SIGNATURE PAGE

This document has been signed electronically on the final page by the following:

	Signatory
Author	PPD Project Role: PPD
Biostatistics Reviewer	PPD Project Role: PPD
Statistical Programming Reviewer	PPD Project Role: PPD
Medical Writing Reviewer	PPD Project Role: PPD

Table of Contents

List of Abbreviations

AE	Adverse Event
AEOI	Adverse Event of Interest
AMR	Alternating Motion Rate
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BAD	Barry Albright Dystonia
BMI	Body Mass Index
BSA	Body Surface Area
CGI-I	Clinician Global Impression of Improvement
CI	Confidence Interval
CSR	Clinical Study Report
C-SSRS	Columbia- Suicide Severity Rating Scale
DDK	Diadochokinetic(s)
eCRF	Electronic Case Report Form
DMC	Data Monitoring Committee
ECG	Electrocardiogram
EQ-5D-3L	EuroQol 5-dimension, 3-level quality of life instrument
EQ-5D-Y	EuroQol-5D Youth Version
EU	European Union
FAS	Full Analysis Set
FIM	Functional Independence Measure
GLM	General Linear Model
ICF	Informed Consent Form
ICH	International Council for Harmonisation
INR	international normalized ratio
ITT	Intent-to-Treat
IWRS	Interactive Web Response System
KM	Kaplan-Meier
LOCF	Last Observation Carry Forward
LSM	Least-Squares Mean
MAR	Missing at Random
max	Maximum
MCAR	Missing Completely at Random
MCMC	Markov Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
min	Minimum
MMRM	Mixed Model for Repeated Measures
MNAR	Missing Not at Random
Neuro-	Quality of life measure for adults/children with neurological disorders

QoL	
OL	Open Label
PK	Pharmacokinetic(s)
PKAN	Pantothenate Kinase Associated Neurodegeneration
PKAN-ADL	Pantothenate Kinase Associated Neurodegeneration Activities of Daily Living Scale
PT	Preferred Term
REML	Restricted maximum likelihood
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Error
SMR	Sequential Motion Rate
SOC	System Organ Class
SUSAR	Suspected Unexpected Serious Adverse Reaction
TEAE	Treatment-Emergent AE
TID	ter in die (three times a day)
TLFs	Tables, Listings and Figures
UPDRS	Unified Parkinson's Disease Rating Scale
US	United States
WeeFIM	Functional Independence Measure for Children
WHO	World Health Organization

1 Introduction

This document summarizes key elements of the statistical analysis plan (SAP) for Protocol 024PKAN15004. The main purpose of this document is to describe the statistical methods to be used for the primary analysis of efficacy and safety endpoints collected during the double-blind portion of the study in support of an interim Clinical Study Report (CSR) for submission to regulatory authorities. This SAP also specifies the scope of analysis for available data collected in the open label portion of the study at the time of the database cut-off for the interim CSR. The plan for analysis of pharmacokinetics (PK) and pharmacogenomics data collected during this study is outside the scope of this SAP.

The SAP is being written in consideration of recommendations outlined in the International Council for Harmonisation (ICH) E9[1] Guideline entitled "Guidance for Industry: Statistical Principles for Clinical Trials" and the most recent ICH E3[2] Guideline, entitled "Guidance for Industry: Structure and Content of Clinical Study Reports."

This SAP is based upon the following study documents:

- Study Protocol, Version 3.0 (February 6, 2019)
- electronic Case Report Form (eCRF), Version 9.0 (March 4, 2019)

2 Study Objectives and Endpoints

The primary objective of this study is to evaluate the efficacy of fosmetpantotenate over 24 weeks in patients with Pantothenate Kinase Associated Neurodegeneration (PKAN).

The secondary objective is to assess the safety and tolerability of fosmetpantotenate in patients with PKAN.

Other objectives are:

- To determine the pharmacokinetics (PK) following multiple doses of fosmetpantotenate in patients with PKAN
- To explore biomarkers of disease, along with their potential response to treatment in patients with PKAN

2.1 Primary Efficacy Endpoint

The Unified Parkinson's Disease Rating Scale (UPDRS) is a comprehensive assessment of the burden and severity of signs and symptoms of parkinsonism captured via systematic interview and neurological examination. It consists of 4 parts that can be used as individual scales or summed for a total score.

Part II of the UPDRS is an in-depth health outcome assessment of motor activities of daily living that are impaired by a wide range of neurological diseases, including PKAN. For this study, Part II of the UPDRS has been adapted and validated for use in PKAN

patients through a systematic, iterative process involving experts and patient advocacy leaders; it is the basis of the primary efficacy endpoint and will be referred to in this SAP as the "PKAN Activities of Daily Living" scale, abbreviated as "PKAN-ADL".

The primary efficacy endpoint is the change in PKAN-ADL total score, from baseline to the end of the 24-week double-blind period.

2.2 Secondary Efficacy Endpoint

The secondary efficacy endpoint is the change in the total score for Part III (motor examination) of the UPDRS from baseline to the end of the 24-week double-blind period.

2.3 Exploratory Efficacy Endpoints

Exploratory efficacy endpoints will include change from baseline to the end of the 24-week double-blind period in the following:

- Clinician Global Impression of Improvement (CGI-I)
- Barry Albright Dystonia (BAD) scale
- Quality of life (QoL):
  - NeuroQoL, upper and lower extremity modules (adult and pediatric versions)
  - EuroQol 5-dimension, 3 level QoL instrument (EQ-5D-3L)/EuroQol-5D Youth Version (EQ-5D-Y)
- Functional independence: Functional Independence Measure (FIM)/Functional Independence Measure for Children (WeeFIM)
- Function measures:
  - Ambulation: 25-foot walk test (in appropriate patients)
  - Speech: Diadochokinetic (DDK) assessments (i.e., alternating motion rate [AMR] and sequential motion rate [SMR])

2.4 Safety Endpoints

The safety and tolerability of fosmetpantotenate will be assessed based on the occurrence of adverse events (AEs) and by serial vital signs including weight, physical examinations, clinical laboratory parameters (chemistry, hematology, coagulation, and urinalysis), the Columbia Suicide Severity Rating Scale (C-SSRS) (in assessable patients), and 12-lead electrocardiograms (ECG).

2.5 Pharmacokinetic Endpoints

Multiple dose PK of fosmetpantotenate in whole blood will be evaluated using a sparse sampling design. Blood samples may also be used for metabolite quantification and/or identification.

2.6 Biomarkers

Exploratory biomarkers will be assessed in plasma. Potential biomarkers assessed may include succinate, lactate, maleate, pyruvate, alanine, acetone, acetoacetic acid, 3-hydroxybutyric acid, 2-hydroxybutyric acid, and acetylCoA. Although other biomarkers may be assessed, the samples will not be used for genetic testing.

2.7 Pharmacogenomics

Pharmacogenomic assessment may be performed on blood samples drawn at screening for drug metabolizing enzymes, transporters, or indicators of therapeutic response or safety. These assessments will only be performed where local regulations permit and where approved by the site's Institutional Review Board/Independent Ethics Committee (IRB/IEC), and will be subject to separate consent/assent by the patient/parent/legal guardian as appropriate.

3 Overall Study Design

This is a pivotal, randomized, double-blind, placebo-controlled, multi-center, 2-arm study evaluating 24 weeks of treatment with fosmetpantotenate or placebo in patients with PKAN aged 6 to 65 years, with a 278 week open-label extension.

Eligible patients will be randomized in a 1:1 ratio to receive either fosmetpantotenate or placebo for a 24-week double-blind period; randomization will be stratified by screening weight and age group (pediatric, adult), as reported by Interactive Web Response System (IWRS). In Protocol Version 1, Pediatric is defined as '<19' and Adult is defined as '≥19.' In subsequent versions of the Protocol, Pediatric is defined as '<18' and Adult is defined as '≥18.'

Treatment in the double-blind period will begin with dose escalation for Days 1-3, followed by the full dose 3 times daily (TID) starting on Day 4. Matching placebo will be given according to the same dose escalation sequence as active investigational product. Study visits during the double-blind period are scheduled at baseline (within 1 to 2 days prior to the start of the double-blind treatment), Days 1 through 4, and Weeks 3, 6, 12, 18, and 24. See Protocol Appendix A (Table 15.1-1) for the schedule of assessments performed during the double-blind period.

After completing the double-blind period, patients will be eligible to receive open-label fosmetpantotenate for up to 278 weeks. During the open-label period, study visits are scheduled at Weeks 27, 30, 36, 48, 60, 72, 84, 96, 108, 120, 146, 172, 198, 224, 250, 276, and 302. See Protocol Appendix A (Table 15.1-2) for the schedule of assessments performed during the open-label period.

In total, patients will be in the study for approximately 302 weeks, with 24 weeks in the randomized, double-blind treatment period and 278 weeks in the open-label period.

4 Statistical Methodology

This section describes the planned methods of analysis for all efficacy and safety endpoints collected during the double-blind and open-label extension periods.

4.1 General Statistical Considerations

Tabulations of summary statistics, graphical presentations, and statistical analyses will be performed using SAS® Version 9.3 or higher. All data collected will be summarized using descriptive statistics, including mean, standard deviation (SD), median, minimum, and maximum for continuous variables, and frequency tables with counts and percentages for categorical variables. Age group categorization will be consistent with the Protocol Version that was applicable at time of patient randomization.

Unless otherwise stated, the "baseline" assessment for the double-blind period will be defined as the last measurement obtained prior to the first dose of blinded investigational product (pre-dose on Day 1). For patients who enter the open-label extension period, the baseline value will be re-defined for the open-label period as the last observation prior to start of open-label fosmetpantotenate, unless otherwise specified.

As appropriate, the scales used as efficacy endpoints to measure neurological impairment, dystonia, motor function, and quality of life will be scored and/or normalized in accordance with user manuals and published scoring procedures for each instrument. For each scale, the prescribed rules for deriving the total score in the presence of 1 or more missing items will be followed.

Unless otherwise noted, if at least 80% of the items for a scale (i.e., PKAN-ADL, UPDRS Part III, BAD, Neuro-QOL) are non-missing, then the prorated total score for the scale is obtained by summing the non-missing item responses, multiplying by the total number of items in the scale, and then dividing by the number of non-missing items.

There are two patients who had a question (item) on the PKAN-ADL mistranslated for screening and baseline visits. In analyses using the Double-Blind Full Analysis Set, the values for this item will be rendered invalid, set to missing, and the PKAN-ADL total score will be prorated as described above. The invalid values will be presented and flagged in the data listing, as originally recorded.

For by-visit summaries, the last non-missing assessment (including repeat assessments, if any) recorded at each visit will be used for summarization. For across visit summaries (e.g. maximum post-baseline value), scheduled, unscheduled (if any) and repeat assessments will all be considered for the value to be analyzed; all observations will be presented in data listings.

Hypothesis-testing will be performed at the two-sided 5% significance level, unless otherwise stated.

4.2 Determination of Sample Size

It is estimated that a total of 74 patients (37 in each treatment group) will provide approximately 80% power to detect a 3-point difference between groups in the average change from baseline score for the PKAN-ADL, based on Student's t-test with alpha=0.05 (two-sided); this calculation assumes a standard deviation of 4.5 points for the change from baseline scores. Allowing for 10% attrition, a total of 82 patients will be randomized to achieve the target sample size of at least 74 patients completing the double-blind period of the study.

4.3 Planned Analysis Schedule

The primary analysis of this study will be performed after all randomized patients have either been assessed at the final visit of the 24-week double-blind treatment period or discontinued treatment and study prior to the final visit of the 24-week double-blind treatment period, and the database is locked/unblinded. The results of the double-blind treatment period, together with available long-term follow-up information in the open-label extension period at the time of the database cut-off, will form the basis of an interim CSR for submission to regulatory authorities.

4.4 Data Sets Analyzed

The following analysis sets will be defined for this study:

Double-Blind Period

Double-Blind Safety Population: The Double-Blind Safety Population will consist of all patients who receive at least 1 dose of blinded investigational product. Patients will be analyzed according to the treatment received. This population will be used for all summaries of patient accountability, demographic and baseline data, and safety information, including AE incidence, for the double-blind period of the study.

Double-Blind Full Analysis Set (FAS): The Double-Blind FAS will consist of all randomized patients with a baseline assessment and at least 1 post-baseline assessment of the primary efficacy endpoint. Patients will be analyzed in the treatment group to which

they were randomized, regardless of treatment actually received. This population will serve as the basis for all efficacy analyses for the double-blind period of the study.

Double-Blind Per-Protocol Set (PPS): The PPS will consist of all FAS patients who complete the double-blind study without any major protocol violations that could impact the interpretation of efficacy. Patients will be analyzed as randomized. This population will be used for sensitivity analyses of the primary and secondary efficacy endpoints (based on PKAN-ADL and UPDRS Part III scores). Specific protocol deviations used to define this analysis set will be finalized prior to database lock.

Open-Label Period

Open-Label Safety Population: The Open-Label Safety Population will consist of all patients who receive at least 1 dose of investigational product after entering the open-label period of the study. This population will be used for all summaries of patient accountability, demographic and Baseline data, and safety information, including AE incidence, for the open-label period of the study.

Open-Label Full Analysis Set (OL-FAS): The OL-FAS will consist of all patients with a baseline assessment and at least 1 post-baseline assessment of the PKAN-ADL in the open-label period of the study. This population will serve as the basis for efficacy analyses from the open-label period of the study.

Fosmetpantotenate Safety Population: The Fosmetpantotenate Safety population will consist of patients who received at least 1 dose of fosmetpantotenate product at any period of the study (double-blind or open-label).

The number of patients in each double-blind analysis set (Safety, FAS, and PPS) will be summarized by treatment group for the double-blind period of the study. For the open-label extension, patients in the Open-Label Safety, OL-FAS, and Fosmetpantotenate Safety Population will be summarized overall and by treatment received in the prior double-blind period. Patients excluded from any analysis set will be listed by reason.

4.5 Disposition of Patients

All patients who meet the inclusion criteria for the study and sign the informed consent form will be accounted for in the disposition table.

The disposition table will present the total number of patients screened for enrollment and the number of patients in each treatment group (fosmetpantotenate and placebo) categorized as follows:

- Randomized
- Evaluated at each study visit
- Completing the double-blind treatment period

- Discontinued from treatment (by reason) for each treatment period
- Discontinued from the study (by reason) for each treatment period

A by-patient data listing of study completion information, including the timing and reason(s) for premature study withdrawal, will be presented.

Protocol deviations will be identified on an ongoing basis by the clinical study team based on the protocol deviation specification document (PDS) and assessed as "minor" or "major" in consultation with the Sponsor. This assessment will be done during the Protocol Deviation Review Meeting, which will take place when pre-determined accrual targets are met and before disclosure of randomization codes. Major deviations from the protocol that may lead to confounding of efficacy results will result in the exclusion of a patient from the per protocol set. The final determination of major protocol deviations and the exclusion of patients from each of the analysis populations will be made prior to database lock.

Unless decided otherwise during the Protocol Deviation Review Meeting, the following will be major protocol deviations:

- Major deviations of inclusion/exclusion criteria;
- Use of prohibited medication or device adjustments post-randomization that may influence the outcome of the study for the patient;
- Participation in another investigational study post-randomization without the prior written authorization of Sponsor or its designee;
- Randomization error (use of wrong medication kit or violation from randomization procedure) during the study;
- <80% or >120% compliance with study drug during the double-blind treatment period;
- Patient met withdrawal criteria but not withdrawn from treatment;
- Unblinding of study medication to blinded staff or patients;
- PKAN-ADL administration error (e.g., administration performed but due to error in translation, the instrument could not be read to the patient at Screening/Baseline Visit; unqualified rater performed the administration of the instrument)

- Failure to report Suspected Unexpected Serious Adverse Reaction (SUSAR) or Serious Adverse Event (SAE) or pregnancy; and

- Informed consent deviation (e.g., improper administration or timing).

Other possible major deviations may be identified during the protocol deviation review meetings.

Protocol deviations will be provided in a listing, sorted by treatment group, site and patient. Major protocol deviations that occur throughout the study will also be summarized by protocol deviation category and treatment group.

4.6 Demographics and Baseline Characteristics

Demographic data (including age, race, ethnicity, sex, height, weight, Body Mass Index (BMI), and Body Surface Area (BSA), pre-treatment clinical characteristics, and baseline disease assessment scores will be summarized by treatment group for the Double-Blind FAS as well as the Double-Blind Safety Population. Descriptive statistics (n, mean, SD, median, minimum, and maximum) will be presented for continuous variables. Frequency distributions (counts and percentages) will be presented for categorical variables. This table will be repeated, stratified by age group (pediatric, adult).

Similar tables summarizing demographic and baseline data will be prepared for patients in the PPS, OL-FAS as well as the Open-Label Safety Population, grouped by prior double-blind treatment assignment and overall. In the event that the FAS and Safety Population are identical for either the double-blind or open-label periods, then the tables will only be produced for the Safety Population.

A by-patient listing will be provided, presenting data collected for demographics and baseline characteristics at Screening, for patients in the Double-Blind Safety Population.

4.7 Medical History

Medical history data obtained at screening will be summarized by MedDRA System Organ Class (SOC), using Version 20.0 of the coding dictionary. These tables will be presented by treatment group for the Double-Blind Safety Population only.

PKAN-Specific Medical History will also be summarized by treatment group using the Double-Blind FAS and the Double-Blind Safety Population. This table will include age of onset, duration of illness (age at screening - age at first symptom), as well as other variables related to the disease history. This table will be repeated for patients in the Open-Label Safety Population.

By-patient listings will be provided, presenting medical history as well as PKAN-specific medical history collected at Screening for all patients in the Double-Blind Safety Population.

4.8 Prior/Concomitant Medications

Prior and concomitant medications, as well as non-drug therapies, will be coded using the World Health Organization (WHO) Drug Dictionary Sep 2018 B3 Format.

4.8.1 Prior and Concomitant Medications

Medications that start and stop prior to the date of first dose of blinded investigational product (placebo or fosmetpantotenate) medication will be classified as Prior only for the double-blind period. If a medication starts before the date of first dose of blinded investigational product and stops on or after the date of first dose of blinded investigational product, then the medication will be classified as both Prior and Concomitant for the double-blind period. Medications that were started or taken after first dose of the double-blinded investigational product and before the first dose of open-label fosmetpantotenate will be classified as Concomitant Only for the double-blind period. Medications will be classified for the open label period of the study in analogous manner.

Medications will be summarized as Prior, Concomitant, and Concomitant only, by treatment group using frequency counts and percentages within anatomical therapeutic chemical (ATC, level 2) and preferred term (PT). Medications that are specifically for PKAN maintenance will also be summarized separately. These summaries will be based on the Double-Blind Safety Population only.

Patients on chronic medications may be counted in a medication category for all periods: prior and concomitant medication for the double-blind period, and prior and concomitant medication for the open-label extension phase.

Although a patient may take more than one medication within the same ATC classification/PT, the patient will be counted only once within an ATC classification/PT for a given study phase. The same patient may contribute to two or more preferred terms in the same classification.

Non-drug therapies will be summarized in a separate table by treatment group using frequency counts and percentages within anatomical therapeutic chemical (ATC, level 2) and preferred term (PT), with no prior/concomitant classification. This summary will be based on the Double-Blind Safety Population.

A by-patient listing will be provided, which presents all medications taken by patients in the randomized population, with flags indicating if the medication is prior, concomitant, etc. for the double-blind or open-label periods. Non-drug therapies will be presented as well, in a separate listing.

4.9 Study Treatment Usage and Compliance

Duration of treatment in weeks [(date of last dose – date of first dose + 1)/7] during the double-blind period and the open-label period will be summarized for the Double-Blind and Open-Label Safety Populations, respectively. For patients in the Fostmetpantotenate Safety Population, duration of treatment following initiation of fosmetpantotenate (double-blind or open-label) will be summarized. The number and percentage of patients with dose changes or interruptions will also be presented.

Compliance is calculated as the number of actual investigational product units (capsules) administered (based on product units returned at each visit) divided by the number of expected investigational product units to be taken during the given treatment period, multiplied by 100%. The number of expected capsules for a given patient is computed by multiplying the patient's daily dosage, in 75mg powder/capsules (for oral suspension), by the daily frequency (e.g. TID) and number of days in the specified dosing period (end date - start date + 1), and then summing these values across all dosing periods within the given treatment period. In the event of a dose interruption, the expected dosage for the given interruption period would be 0 capsules. Compliance will be summarized descriptively for each treatment group, including frequency tables to indicate the number of patients who took between 80% and 120% of the doses prescribed for each treatment period.

A by-patient listing will be provided, which will present total drug accountability, along with information on compliance and dose modifications by treatment period.

4.10 Efficacy Analysis

Analysis of the efficacy endpoints for the Interim Analysis will be limited to data collected during the double-blind period. Open-Label efficacy data will be included in the listings only. Following the Interim Analysis, additional analyses including open-label efficacy data will be described in a later version of the SAP, for presentation as part of the Final Analysis.

4.10.1 Primary Efficacy Endpoint

The PKAN-ADL consists of 12 items related to activities of daily living, including eating, dressing, and walking. Each item has responses ranging from 0-4, with a higher value indicating greater disability in the given activity. To compute the total score, responses are summed across the 12 items.

For analysis purposes, data from the pediatric and adult age groups will be pooled. All descriptive summary tables will be repeated using age group as a subgroup.

Descriptive statistics will be presented for the PKAN-ADL total score, with absolute and percent change from baseline, by treatment group for each visit of the double-blind

period. For post-baseline assessments, categories for change from baseline total score (e.g., >0, =0, <0, ≤ -1, ≤ -2, ≤ -3, ≤ -4), categories for percent improvement (i.e., decrease) in total score from baseline (i.e., <0%, =0% >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), and categories for total score status (improvement, no change, or worsening) will be presented. These categories may be adjusted in a way that ensures that the categories contain at least 10% of patients. This will be run using the Double-Blind FAS, and repeated using the PPS, as a sensitivity analysis.

Change from baseline will be calculated by subtracting the baseline total score from the post baseline total score. Percent change from baseline will be calculated as:

100(Observed Score at Post-Baseline Visit – Observed Score at Baseline) / Observed Score at Baseline*

Furthermore, a chi-squared test will be performed at each post-baseline visit to determine treatment differences between the percentages of patients who had a total score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

Frequency tables will display change from baseline for each of the 12 items, as well as shift from baseline, by treatment group for each post-baseline assessment of the double-blind period. In addition, categories for change from baseline item score (i.e., >0, =0, <0, ≤ -1, ≤ -2, ≤ -3), and categories for item score status (improvement, no change, or worsening) will be presented. These summaries will be run using the Double-Blind FAS and repeated using the PPS, as a sensitivity analysis.

Figures will be presented which show the PKAN-ADL total score for each visit during the double-blind period, as well as change from baseline score, by treatment group. This will be run using the Double-Blind FAS and repeated using the PPS.

In addition, a by-patient listing will be provided, which presents PKAN-ADL scores (total and for each item) across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

The change from baseline in PKAN-ADL total scores will be used to test the following hypotheses:

Null hypothesis:

There is no difference between the fosmetpantotenate and placebo groups, as measured by the change in PKAN-ADL total score from baseline to the end of the 24-week double-blind treatment period, with adjustment for age group and baseline PKAN-ADL total score.

Alternate hypothesis:

There is a difference between the fosmetpantotenate and placebo groups, as measured by the change in PKAN-ADL total score from baseline to the end of the 24 week double-blind treatment period, with adjustment for age group and baseline PKAN-ADL total score.

That is:

> H_0: Change from baseline PKAN-ADL Total Score at Week 24 $_{(\text{fosmetpantotenate})}$ = Change from baseline PKAN-ADL Total Score at Week 24 $_{(\text{Placebo})}$ *estimated via repeated measure analysis, with adjustment for age group and baseline score*
>
> H_1: Change from baseline PKAN-ADL Total Score at Week 24 $_{(\text{fosmetpantotenate})} \neq$ Change from baseline PKAN-ADL Total Score at Week 24 $_{(\text{Placebo})}$ *estimated via repeated measure analysis, with adjustment for age group and baseline score*

The primary efficacy analysis will be based on the FAS population in the double-blind period of the study. For each patient, the change from baseline in PKAN-ADL total scores at Weeks 3, 6, 12, 18, and 24 of the double-blind period will be used for analysis. The treatment effect will be evaluated on the basis of a mixed model for repeated measures (MMRM) analysis to assess data from all visits simultaneously. The model will include fixed effects for treatment, age group (pediatric, adult), visit and treatment by-visit interaction, with visit as the repeating factor and patient as a random effect, and baseline PKAN-ADL total score as a covariate. The PKAN-ADL total score obtained prior to the first dose of investigational product will serve as each patient's baseline value.

Patients with PKAN-ADL administration errors will be included in the Double-Blind FAS, but excluded from the PPS, as per Section 4.5. Imputation rules for the data affected by the translation issue can be found in Section 4.1.

The linear mixed-effect model will be implemented using PROC MIXED in SAS with the restricted maximum likelihood (REML) estimation method. An unstructured variance-covariance matrix will be used to model the within-patient errors. [**Note**: In the unlikely event that the computational algorithm fails to converge, the following structures will be executed in the order specified (essentially in decreasing order of complexity) until convergence is achieved: Toeplitz with heterogeneity, autoregressive with heterogeneity by Week, compound symmetry with heterogeneous variances by Week, autoregressive, Toeplitz, and compound symmetry without heterogeneous variances by Week.]

Significance tests will be based on least squares means (LSMs) using Type III sum of squares. The Kenward and Roger method will be used for calculating the denominator degrees of freedom for tests of fixed effects.

The PROC MIXED SAS coding is as follows:

```
proc mixed data=DATA;
   class TRT WEEK PATIENT;
   model C_score = BASE_score AGE TRT WEEK TRT*WEEK / solution
ddfm=kr;
   repeated WEEK / subject=PATIENT type=un;

run;

where:
 TRT = treatment group (0 = placebo, 1 = fosmetpantotenate)
 WEEK = week 3, 6, 12, 18, 24
 AGE = age group (0=Pediatric, 1= Adult)
 C_score = change in score from baseline
 BASE_score = baseline score
```

The LSMs for each treatment, the difference between the LSMs, and the standard error (SE) and 95% confidence interval (CI) associated with the difference will be reported at each study visit. The primary treatment comparison will be based on the test of contrast between treatment effects using the LSM estimate at Week 24, adjusted for baseline score and age group from the Type III analysis.

To estimate the mean change from baseline for PKAN-ADL total scores in each treatment group (placebo or fosmetpantotenate) at Week *i* (*i* = 3, 6, 12, 18, and 24), the following estimate statement (as an example) will be utilized:

```
estimate 'Placebo Group at Week 24'
    intercept   1
    AGE         <proportion in age group=0   proportion in age group=1>
    BASE_score  <mean of BASE_score>
    TRT         1 0
    WEEK        0 0 0 0 1
    TRT*WEEK    0 0 0 0 1
                0 0 0 0 0 / cl;
estimate 'FOSMETPANTOTENATE   Group at Week 24'
    intercept   1
    Age         <proportion in age group=0   proportion in age group=1>
    BASE_score  <mean of BASE_score>
    TRT         0 1
    WEEK        0 0 0 0 1
    TRT*WEEK    0 0 0 0 0
                0 0 0 0 1  / cl;
```

To estimate the treatment effect (i.e., the difference between fosmetpantotenate and placebo) at Week *i* (*i* = 3, 6, 8, 12, and 24), the following estimate statement (as an

example) will be utilized:

```
estimate 'Treatment Effect at Week 24'
    TRT          -1 1
    TRT*WEEK      0 0 0 0 -1
                  0 0 0 0 1  / cl;
```

The point estimates and corresponding 95% CIs for the treatment effect at each week will be presented for the FAS (double-blind period).

4.10.1.1 Sensitivity Analysis

To check that primary efficacy findings are robust, sensitivity analyses will be performed, as follows:

- The MMRM described in section 4.10.1 will be repeated using the PPS, and presented in a similar manner. Of note, patients who had a PKAN-ADL administration error will be excluded from the PPS.
- In addition, analysis of covariance (ANCOVA) models will be applied to change from baseline in PKAN-ADL total scores assessed after 3, 6, 12, 18, and 24 weeks of treatment. For each week analyzed, the model will include treatment and age group (pediatric, adult) as fixed effects, and baseline PKAN-ADL total score as a covariate.

  The ANCOVA analysis will be performed for the Double-Blind FAS using PROC GLM in SAS. The models will be run using the observed case (OC) and last-observation-carried-forward (LOCF) procedure to account for incomplete data. In each case, estimates of adjusted treatment means and the adjusted mean difference between treatment groups will be presented at each time point, along with the two-sided 95% CIs and p-values for treatment comparisons.

  The LOCF procedure will be applied for post-baseline observations of the PKAN-ADL total score in which more than 20% of the items are missing and cannot be imputed in the manner described in Section 4.1. The last observed post-baseline PKAN-ADL total score, including those from unscheduled visits, will be imputed to replace the missing value of PKAN-ADL total score.

  The PROC GLM SAS coding, to be run for each visit and each OC/LOCF procedure, is as follows:

```
proc glm data=DATA;
 class TRT AGE;
 model week_x_total=base_score TRT AGE / solution;
 lsmeans TRT/ pdiff stderr;
```

```
run;
```

4.10.1.2 Handling of Dropouts and/or Missing Data

The MMRM approach used for the primary efficacy analysis accounts for data "missing at random" through a variance covariance structure. The ANCOVA models (described above) employ both the OC and LOCF methods to explore the impact of drop-outs and missing data. However, mixed models have been shown to be less biased under Missing at Random (MAR) and Missing Completely at Random (MCAR) conditions compared with LOCF or Completer analyses (Prakash, 2008).

To assess the appropriateness of the MAR assumption, additional sensitivity analyses using a tipping point approach will be used to assess how extreme and detrimental outcomes among patients with missing data must be to overwhelm the treatment effect attained in those patients who had complete data. In these analyses, the same basic analytic model will be used as in the primary analysis, but a multiple imputation approach will be implemented to allow for varying impact of missing data by incorporating a shift parameter in the imputation model. The range of values for the shift parameter will explore the varying missing data assumptions, from MAR to Missing Not at Random (MNAR), including the scenario where the imputation model is completely based on the control group (i.e., control-based imputation).

Following the approach described in Ouyang (2017), the following imputation procedures will be implemented:

1. Intermittent missing values before a discontinuation event (i.e., discontinuation of randomized therapy or early permanent dropout): The MAR assumption is reasonable for intermittent missing values before a discontinuation event since the values of the endpoint before and after the intermittent missing value are known. Imputation of intermittent missing data will be accomplished using the Markov Chain Montel Carlo (MCMC) option in SAS PROC MI by treatment group to impute intermittent missing data without a monotone missing pattern, prior to performing imputation of values following the discontinuation event.

2. Missing data following a discontinuation event (i.e., discontinuation of randomized therapy or early permanent dropout): Imputation of the post-discontinuation missing data by treatment group under the assumption of MAR using the regression option from the monotone statement of SAS PROC MI. Baseline and post-baseline scheduled visits will be used in the regression option to impute the missing values.

3. Steps 1 and 2 will be repeated to generate 30 multiply-imputed datasets.

4. For the fosmetpantotenate-treated patients, the post-discontinuation imputed values from step 3 will be made worse by subtracting a delta defined as k=100% times the treatment differences (means) obtained for the post-discontinuation visits obtained from the specified mixed model analysis (i.e., the analytic model) of the endpoint. No adjustment from the MAR assumption will be implemented for the placebo-treated patients (i.e., k = 0%).

5. The mixed model (i.e., the analytic model) analysis described for the endpoint will be performed on each set of the newly "complete" data (observed data plus imputed values).

6. Using Rubin's approach (Rubin 1987), the estimated treatment effects are combined across imputations. This will be accomplished using SAS PROC MIANALYZE.

7. Repeat steps 4 through 6 with increasing values of k (e.g., 120%, 140%…200%…) until the significance of the pre-specified analysis assuming MAR is overturned (e.g., from p-value <0.05 to p-value ≥0.05). The worsening by k times the treatment differences is applied to all the post discontinuation imputed values for each set of "complete" data from step 3. As an example, if a fosmetpantotenate patient had 2 post-discontinuation visits imputed, the values of both of these 2 imputed visits would be made worse by subtracting k times the fosmetpantotenate versus placebo treatment differences at these 2 visits.

The iterations for each value of delta will be summarized in a table where the estimated treatment differences for the treatment comparisons will be presented together with 95% confidence intervals for the differences and p-values for the treatment comparisons.

4.10.1.3 Exploratory Subgroup Analysis

The influence of potential prognostic factors on the primary efficacy endpoint will be assessed via the mixed model described for the primary analysis, using the Double-Blind FAS. That is, the mixed model will include baseline PKAN-ADL total score and additional prognostic factors (each considered univariately) as covariates, with treatment group, Week, treatment*Week, and treatment*prognostic factor as fixed effects. The factors to be assessed are as follows:

- Age at screening (pediatric, adult)
- Age at disease onset (<median, ≥ median)
- Sex (male, female)
- Baseline PKAN-ADL total score (1st quartile, 2nd quartile, 3rd quartile, 4th quartile).
- Duration of Illness (<median, ≥ median)

Factors found to have a significant effect on response ($p < 0.10$) will be examined further, to understand the relationship between prognostic factor and response. For all factors, the

LSMs for each treatment, the difference between the LSMs, and the standard error (SE) and 95% confidence interval (CI) associated with the difference will be reported at each study visit, using the MMRM which includes the given factor as a covariate.

4.10.2 Secondary Efficacy Endpoint

Part III of the UPDRS consists of 27 items, which correspond to 14 domains related to motor abilities such as tremor, stability, and bradykinesia. Each item has responses ranging from 0-4, with a higher value indicating greater disability. To compute the UPDRS Part III total score, responses are summed across the 27 items, and accordingly, will range from 0-108. For domain totals, responses will be summed across all of the items in a given domain (when domain corresponds to multiple items).

For analysis purposes, data from the pediatric and adult age groups will be pooled. All descriptive summary tables will be repeated using age group as a subgroup.

Descriptive statistics will be presented for the UPDRS Part III total score, with absolute and percent change from baseline, by treatment group for each visit of the double-blind period. For post-baseline assessments, categories for change from baseline UPDRS Part III total score (e.g., >0, =0, <0, ≤ -1, ≤ -2, ≤ -3, ≤ -4), categories for percent improvement (i.e., decrease) in UPDRS Part III total score from baseline (i.e., <0%, =0%, >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), and categories for UPDRS Part III total score status (improvement, no change, or worsening) will be presented. These categories may be adjusted in a way that ensures that the categories contain at least 10% of patients. This will be run using the Double-Blind FAS and repeated using the PPS, as a sensitivity analysis.

Frequency tables will display change from baseline for each of the 14 domains, by treatment group for each visit of the double-blind period, using categories for change from baseline domain score (i.e., >0, =0, <0, ≤ -1, ≤ -2, ≤ -3), and categories for domain score status (improvement, no change, or worsening). These summaries will be run using the Double-Blind FAS and repeated using the PPS, as a sensitivity analysis.

The MMRM model described above for the primary efficacy endpoint will also be performed for the secondary efficacy endpoint: change from baseline UPDRS Part III total score. The LSMs for each treatment, the difference between the LSMs, and the standard error (SE) and 95% confidence interval (CI) associated with the difference will be reported at each study visit. The analysis will be based on the Double-Blind FAS, and will be repeated using the PPS, as a sensitivity analysis.

The ANCOVA models, as described in section 4.10.1.1, will be run for change from baseline in UPDRS Part III total scores assessed after 3, 6, 12, 18, and 24 weeks of treatment for patients in the Double-Blind Full Analysis Set. The models will be run

using the observed case (OC) and last-observation-carried-forward (LOCF) procedure to account for incomplete data.

The exploratory subgroup analysis, as described in section 4.10.1.3, will also be run for the UPDRS Part III total score.

Furthermore, a chi-squared test will be performed at each post-baseline visit to determine treatment differences between the percentages of patients who had a total score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

Figures will be presented which show the UPDRS Part III total score for each visit during the double-blind period, as well as change from baseline score, by treatment group. This will be run using the Double-Blind FAS and repeated using the PPS.

A by-patient listing will be provided, which presents UPDRS Part III scores (total and for each of the 27 items) across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

4.10.3 Exploratory Efficacy Endpoints

Exploratory efficacy endpoints that are continuous rating scales recorded at multiple post-baseline visits (i.e., BAD Scale, Neuro-QOL, Euro-QOL, 25-foot walk test) will be evaluated using the MMRM approach described above for the primary efficacy endpoint. For endpoints obtained only at the baseline and final (Week 24) visits of the double-blind period (i.e., FIM/WeeFIM, DDK), the ANCOVA model will be used. The model will be run using the observed case (OC) procedure only. Summary statistics for efficacy measures will include graphical plots of mean and mean change from baseline values at each study visit during the double-blind period, by treatment group.

For analysis purposes, data from the pediatric and adult age groups will be pooled. All descriptive summary tables will be repeated using age group as a subgroup.

Treatment effects on categorical outcomes will be examined using frequency tables and tested using the Cochran-Mantel-Haenszel (CMH) chi-square methods with stratification by age group (pediatric, adult) (Mantel and Haenszel, 1959).

For the Interim Analysis, the Double-Blind FAS only will be used for analyses involving exploratory efficacy endpoints.

Statistical methods and estimation procedures are described below for each of the exploratory endpoints.

4.10.3.1 Clinician Global Impression of Improvement (CGI-I)

The CGI-I is measured on a 7-point scale, ranging from 1=very much improved to 7=very much worse since the start of treatment. Frequency tables will display the distribution of CGI ratings in each treatment group at Weeks 3, 6, 12, 18, and 24 of the double-blind treatment period. In addition, the percentages of patients who had a score indicating improvement, no change, or worsening, will be presented. This will be repeated using age group (pediatric, adult) as a subgroup.

The Cochran-Mantel-Haenszel test (CMH) row-mean score test (van Elteren test) will be used to evaluate the difference between treatment groups in average CGI-I ratings at each visit during the double-blind period, with stratification by age group (pediatric, adult).

A chi-squared test will be performed at Week 24 to determine treatment differences between the percentages of patients who had a score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

A by-patient listing will be provided, which presents CGI-I scores across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

4.10.3.2 Barry Albright Dystonia (BAD) scale

The BAD scale assesses the signs of dystonia in 8 regions (eyes, mouth, neck, trunk, upper left/upper right extremities, and lower left/right extremities), each rated on a 5-point ordinal scale ranging from 0=absent to 4=severe. Thus, the total BAD score ranges from 0 to 32, with higher scores denoting greater severity of dystonia.

Frequency tables will display change from baseline for each of the 8 regions, as well as shift from baseline, by treatment group across each visit of the double-blind period, using categories for change from baseline region score (i.e., >0, $=0$, <0, ≤ -1, ≤ -2, ≤ -3), and categories for region score status (improvement, no change, or worsening).

Summary statistics will be presented for the total BAD score, including absolute and percent change from baseline, across each visit of the double-blind period. In addition, categories for change from baseline total BAD score (e.g., >0, $=0$, <0, ≤ -1, ≤ -2, ≤ -3, ≤ -4), categories for percent improvement (i.e., decrease) in total BAD score from baseline (i.e., $<0\%$, $=0\%$, $>0\%$, $\geq 10\%$, $\geq 20\%$, $\geq 30\%$, $\geq 40\%$, $\geq 50\%$), and categories for total BAD score status (improvement, no change, or worsening) will be presented. These tables will be repeated using age group (pediatric, adult) as a subgroup.

Furthermore, a chi-squared test will be performed at each post-baseline assessment to determine treatment differences between the percentages of patients who had a total score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

The treatment effect on change from baseline in total BAD score will be evaluated at Week 12 and the final (Week 24) visit using an MMRM, similar to the one described for the primary endpoint. The model will be run using the Double-Blind FAS.

Figures will be presented which show the total BAD score for each visit during the double-blind period, as well as change from baseline score, by treatment group.

A by-patient listing will be provided, which presents BAD scores (total and for each region) across all assessments in the double-blind and open-label periods for the Double-Blind FAS.

4.10.3.3 Quality of Life (QoL)

4.10.3.3.1 Neuro-QoL

The Neuro-QoL scale includes an upper extremity module (fine motor and activities of daily living) and a lower extremity module (mobility) with 19 items (20 items in the Pediatric Version).

Frequency tables will display change from baseline scores for each item of each module, including shift from baseline, by treatment group for Week 12 and Week 24. In addition, categories for change from baseline item score (i.e., <0, $=0$, >0, ≥ 1, ≥ 2, ≥ 3, ≥ 4), and categories for item score status (improvement, no change, or worsening) will be presented. For these tables, data from the pediatric and adult versions will be presented in separate outputs.

For analytic purposes, responses across each module will be summed, and if necessary, rescaled, such that the total score will range from 0-100 for each module in both the Pediatric and Adult Versions. This methodology is described below.

For the Adult Version of the Neuro-QOL Upper Extremity module, in which there are 20 items with each response ranging from 1-5, the total score is computed via summation, and no rescaling is necessary. In the Adult Version of the Neuro-QOL Lower Extremity module, there are 19 items with each response ranging from 1-5. The total score is thus computed by summing the responses across the 19 items, and then multiplying this sum by (20 / 19).

In each of the Pediatric Version of the Neuro-QOL Upper and Lower Extremity modules, there are 20 items with each response ranging from 0-4. For all items, each response is rescaled to correspond to a 1-5 scale, and then summed to produce the total score.

Within each Neuro-QOL module, the rescaled total scores from the adult and pediatric versions of the assessment will be pooled. Descriptive statistics of the rescaled total score for each of the Neuro-QOL modules, as well as absolute and percent change from baseline, will be presented for patients in the Double-Blind FAS at each post-baseline assessment during the double-blind period. In addition, categories for change from baseline rescaled total score (e.g., <0, $=0$, >0, ≥ 1, ≥ 2, ≥ 3, ≥ 4), categories for percent

improvement (i.e. increase) in rescaled total score from baseline (i.e., <0%, =0% >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), categories for rescaled total score status (improvement, no change, or worsening) will be presented for each module. This table will be repeated using age group (pediatric, adult) as a subgroup.

Chi-squared tests will be performed for each module at each post-baseline assessment to determine treatment differences between the percentages of patients who had a re-scaled total score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

In addition, the change in Neuro-QoL assessments (rescaled score for each module) from baseline to the Week 12 and final (Week 24) visit will be evaluated at each time point using an MMRM, similar to the one described for the primary endpoint. The model will be run using the Double-Blind FAS.

A by-patient listing will be provided, which presents the raw responses to all items in the Neuro-QOL assessment, along with the rescaled total score for each module, across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

4.10.3.3.2 Euro-Qol

The EQ-5D-3L (for adult patients) measures 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each rated at 3 levels: 1=no problems, 2=some problems, 3=extreme problems. The scale also records the patient's self-rated health on a visual analogue scale (VAS) from 0=the worst health you can imagine, to 100=the best health you can imagine. The EQ5DY (for pediatric patients) measures the same 5 dimensions on a similar 3-point scale and uses the same VAS for the patient to self-rate health.

Frequency counts of the change from baseline score across each of the 5 dimensions, as well as shift from baseline, will be presented for each visit during the double-blind period. Categories for change from baseline dimension score (i.e., >0, =0, <0, ≤ -1), and categories for dimension score status (improvement, no change, or worsening) will be presented at Week 24, as well. These tables will be repeated using age group (pediatric, adult) as a subgroup.

Furthermore, each patient's 5-digit EQ-5D state will be transformed into a single EQ-5D™ index score by following the algorithm, or value set, specified for his or her relevant country. These value sets are derived based off surveys conducted in representative samples of the given countries. The value set is anchored at 0=dead and 1=full health. The EURO-QOL group provides the value set for performing these calculations, and one such reference table can be found in Janssen (2014). If there are any

missing data across the 5 dimensions at a given visit, then the index score cannot be validly calculated and is considered missing.

Descriptive statistics of the EQ-5D™ index score will be presented for patients in the Double-Blind FAS at each visit of the double blind period, including absolute change from baseline. Categories for percent improvement (i.e., increase) in EQ-5D™ index score from baseline (i.e., <0%, =0% >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), and categories EQ-5D™ index score status (improvement, no change, or worsening) will be presented. In addition, descriptive statistics of the patient's self-rated health on the VAS, as well as absolute change from baseline, will be presented for each visit of the double-blind period. This table will be repeated using age group (pediatric, adult) as a subgroup.

A chi-squared test will be performed at each post-baseline assessment to determine treatment differences between the percentages of patients who had an EQ-5D™ index score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

Furthermore, the change in each Euro-QoL assessment (EQ-5D™ index score, and the EQ VAS) from baseline to the Week 12 and final (Week 24) visit will be evaluated using an MMRM, similar to the one described for the primary endpoint. The models will be run using the Double-Blind FAS.

A by-patient listing will be provided, which presents Euro-QOL scores (score on 5 dimensions, EQ-5D™ index score, and VAS score) across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

4.10.3.4 Functional Independence Measure

The FIM (or WeeFIM for pediatric patients) assesses motor tasks (13 items) and cognitive tasks (5 items), each rated on a 7-point ordinal scale, with higher values denoting greater functional independence. The total score ranges from 18 (total assistance) to 126 (complete independence); the total subscores range from 13 to 91 for the motor component, and from 5 to 35 for the cognitive component. In the WeeFIM, the motor component is comprised of two domains: a self-care domain with 8 items (subscores ranging from 8 to 56) and a mobility domain with 5 items (subscores ranging from 5 to 35). If an item is missing, then the total score and corresponding subscore will not be computed.

A descriptive summary of the absolute change from baseline in motor and cognitive components of the scale, along with the total score, will be presented by treatment group, overall (pooled across age groups) and within each age group (pediatric, adult). For the pediatric group, the subscores for the self-care and mobility domains will also be summarized. In addition, at Week 24, categories for change from baseline FIM total score (e.g., <0, =0, >0, ≥ 1, ≥ 2, ≥ 3, ≥ 4), categories for percent improvement (i.e., increase) in

FIM total score from baseline (i.e., <0%, =0% >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), and categories for FIM total score status (improvement, no change, or worsening) will be presented.

Frequency counts of the change from baseline score across each FIM item, as well as shift from baseline, will be presented for Week 24. Categories for change from baseline FIM item score (i.e., <0, =0, >0, ≥ 1, ≥ 2, ≥ 3, ≥ 4), and categories for FIM item score status (improvement, no change, or worsening) will be presented, as well. In addition, frequency counts will be presented by treatment group for patient locomotion (walk, wheelchair or both), comprehension (auditory, visual, or both) and expression (vocal, nonvocal, or both) will be presented at baseline and Week 24.

A chi-squared test will be performed at Week 24 to determine treatment differences between the percentages of patients who had a total score indicating improvement, no change, or worsening. Proportions, with estimated 95% CI, will be reported for each treatment group, along with a p-value corresponding to the treatment comparison.

The change in FIM (or WeeFIM) total score from baseline to the Week 24 visit will be evaluated using an ANCOVA model, with treatment and age group (pediatric, adult) as fixed effects, and baseline score as a covariate, pooling data across age group (pediatric, adult). The model will be run using the observed case. The adjusted mean treatment difference will be reported together with its 95% confidence interval.

A by-patient listing will be provided, which presents FIM scores (total and for each component) across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

4.10.3.5 Function Measures

Ambulation: Patients able to attempt this assessment will have 2 trials to walk 25 feet, and the average time (in seconds) of these trials will be taken. Patients must successfully complete 2 trials in order for the assessment to have been considered completed. If there is only data available for 1 trial, for reasons other than a patient's inability to perform the test, then the time for the single, available trial will be considered as the average.

Descriptive statistics will be used to summarize average walking speeds (in seconds) at each study visit of the double-blind treatment period (Baseline, Weeks 6, 12, 18, and 24) for patients in the Double-Blind FAS, along with absolute and percent change from baseline for post-baseline visits. Categories for percent improvement (i.e., decrease) in average walking speed (in seconds) from baseline (i.e., <0%, =0% >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), and categories for average walking speed (improvement, no change, or worsening) will be presented, as well. This table will be repeated using age group (pediatric, adult) as a subgroup.

Moreover, the walking speeds will be converted to z-scores, using the baseline values of patients in the Double-Blind FAS as reference values. The z-score conversion will follow the below formula:

(Observed value [in seconds] – average of all baseline values [in seconds]) / standard deviation of all baseline values [in seconds].

Then, the z-score will be adjusted (multiply z-score by -1) so that a higher z-score will correspond to a better speed (fewer seconds). For patients who are unable to complete the 25-foot walk test, a z-score will be imputed using the worst observed z-score across all patient visits.

Descriptive statistics of the z-score, including absolute and percent change from baseline, will be presented for patients in the Double-Blind FAS at each visit during the double-blind period. Depending on the number of patients with available data, the MMRM model (as described for the primary endpoint) will be used to evaluate the treatment effect on change from baseline walking speed z-score. The difference between adjusted treatment means will be reported, together with its 95% CI.

Frequency counts will be presented showing percentages of patients using given assistive devices at each visit, percentages of patients who failed to complete 1 or more trials, and percentages of patients in the Double-Blind FAS who reported a change in level of assistance needed.

In addition, figures will be presented which show the walking speed for each visit during the double-blind period, as well as change from baseline speed by treatment group. A by-patient listing will be provided, which presents the results from the 25-foot walk test (including Trials 3 and 4, if necessary) across all assessments in the double-blind and open-label periods for patients in the Double-Blind FAS.

Speech: Diadochokinetic test (DDK) assessments will provide 4 scores for analysis: three scores for the alternating motion rate (AMR) and one for the sequential motion rate or (SMR). For the three AMR scores, patients will be required to recite a single syllable (“puh”, “tuh”, and “kuh”) as rapidly as possible in one breath. The AMR will be derived as the number of repetitions over the duration (in seconds). For the SMR scores, patients will be required to recite multiple syllables (“puhtuhkuh”) as rapidly as possible in one breath. The SMR will be derived as the number of repetitions over the duration (in seconds).

The AMR and SMR will be analyzed using descriptive statistics at baseline and Week 24 of the double-blind period, including absolute and percent change from baseline. Categories for percent improvement (i.e., decrease) in AMR/SMR (i.e., <0%, =0% >0%, ≥10%, ≥20%, ≥30%, ≥40%, ≥50%), and categories for AMR/SMR status (improvement, no change, or worsening) will be presented, as well. This will be repeated using age group (pediatric, adult) as a subgroup.

Moreover, for each score, the change from baseline to the Week 24 visit of the double-blind period will be evaluated using an ANCOVA model, with treatment and age group (pediatric, adult) as fixed effects, and baseline score as a covariate. The models will be run using observed cases. The adjusted mean treatment difference will be reported together with its 95% confidence interval.

For each of the 4 repetition tasks ("puh", "tuh", "kuh", "puhtuhkuh"), the number of repetitions performed will also be summarized descriptively at baseline and Week 24 during the double-blind period. This will be repeated using age group (pediatric, adult) as a subgroup.

A by-patient listing will be provided, which presents the duration, repetitions, and repetition rates for the AMR and SMR across all assessments in the double-blind and open-label periods for the Double-Blind FAS.

4.10.4 Adjustment for Multiple Hypothesis Testing

To control the Type 1 familywise rate at 5%, the serial gate-keeping approach (proposed by Westfall and Krishen, 2001) will be used to test the primary and secondary efficacy endpoints (ordered by clinical relevance and importance) during the double-blind period in the above specified sequence. It will be implemented in the following way: If (and only if) the null hypothesis associated with the primary endpoint is rejected at the 0.05 significance level, then the secondary efficacy endpoint will be tested at the 0.05 significance level. The plan for analysis of other endpoints during the double-blind period is purely exploratory, so no further adjustment is warranted. Similarly, efficacy analyses of endpoints collected during the open-label period are also exploratory, and do not warrant adjustment.

4.10.5 Safety Data Monitoring

An independent Data Monitoring Committee (DMC) will be established to review safety data periodically throughout the trial.

The DMC will operate under a charter that is approved by all committee members and the Sponsor. The scope of the review and procedures in place to maintain the integrity of the ongoing study will be described in the charter.

4.11 Safety Analysis

For the Interim Analysis, the safety analyses will cover all data collected until the database cut-off, including data from the open-label period. Safety analyses will be performed using data from the Double-Blind and Open-Label Safety Populations. Tables using the Double-Blind Safety Population will include only data collected during the

double-blind period. Tables using the Open-Label Safety Population will include only data collected during the open-label period.

In addition, select safety analyses will be performed using the Fosmetpantotenate Safety Population. Tables using the Fosmetpantotenate Safety Population will include data collected following the first exposure for fosmetpantotenate (i.e., double-blind and open-label data for patients administered fosmetpantotenate at baseline of double-blind period, open-label data only for patients first administered fosmetpantotenate at baseline of the open-label period.)

Safety of fosmetpantotenate will be determined on the basis of treatment-emergent AEs (TEAEs), vital signs, weight, physical examinations, clinical laboratory assessments, the Columbia Suicide Severity Rating Scale (C-SSRS) (in assessable patients), and ECG findings.

For the Double-Blind Safety Population, table summaries and figures will be presented by treatment group as appropriate; for the Open-Label and Fosmetpantotenate Safety Populations, safety summaries will be presented for the entire cohort, and (when indicated) by treatment received during the prior double-blind period.

4.11.1 Adverse Events (AEs)

All AEs will be coded by Preferred Term and System Organ Class using MedDRA coding dictionary Version 20.0.

TEAEs in the double-blind period are defined as AEs that are new or are a worsening of an existing condition that begin from the day of first dose of investigational product until the day after the last dose for the double-blind treatment period. TEAEs in the open-label period are defined as AEs that are new or are a worsening of an existing condition that begin from the day of first dose of open-label investigational product. TEAEs in the fosmetpantotenate safety population are defined as AEs that are new or are a worsening of an existing condition that begin from the day of first dose of (blinded or open-label) fosmetpantotenate.

An overview of adverse events will be provided for each of the three Safety Populations that summarizes incidence in the following categories, overall, by age group (pediatric, adult), and by sex:

- Any AEs
- Any TEAEs
- TEAEs by relationship to study medication (Not related, Unlikely Related, Possibly Related or Related)
- TEAEs by severity (mild, moderate, severe)

- Any TEAEs with an outcome of death
- TESAEs
- Treatment discontinuation due to TEAEs

The incidence of TEAEs will be summarized for each treatment group by Preferred Term as well as by Preferred Term within System Organ Class. The incidence of TEAEs by severity and relationship to treatment will be summarized for each treatment group by System Organ Class and Preferred Term. A patient with several occurrences of the same AE will be counted once and classified by the most severe event (or the event with the closest relationship to study drug). If severity or causality is missing, the worst case will be assumed (i.e., "Severe" if severity is missing and "Drug Related" if causality is missing.)

In addition, the incidence of Adverse Events of Interest (AEOIs), TESAEs, TEAEs leading to treatment discontinuation, and TEAEs leading to death will be summarized for each treatment group by System Organ Class and Preferred Term.

Incidence calculations will be based on the numbers and percentages of patients with AEs, and sorted by decreasing frequency in the fosmetpantotenate group. Although a MedDRA Preferred Term may be reported more than once for a patient, that patient will be counted only once for that MedDRA Preferred Term. For these summaries, patients with multiple AEs will be counted only once per SOC and PT. Furthermore, for each category within these summaries, the number of adverse events will be presented.

The above incidence tables will each be run using the Double-Blind, Open-Label, and Fosmetpantotenate Safety Populations.

In addition, time to event analyses (first event) will be performed for the following adverse event categories for patients in the Double-Blind Safety Population: dystonia, any AE with gastrointestinal disorder as SOC, any AEOI. Time to event endpoints will be summarized using the Kaplan-Meier method; median event times and 2-sided 95% confidence interval for each median will be provided (Brookmeyer, 1982). Time to event data will be censored using the date of a patient's last visit during the double-blind period. The event or censoring time (days) will be calculated as:

$$\textit{Date of event/censoring} - \textit{Date of randomization} + 1$$

Moreover, for each AE category, a log-rank test will be performed to compare time to event between the treatment groups. The Cox proportional hazard (PH) model will be fitted and the estimated hazard ratio and 2-sided 95% confidence interval will be provided. Finally, estimates of the time to event curves obtained from the Kaplan-Meier method will be presented and displayed graphically, spanning the double-blind period.

For the above three adverse event categories, the duration of each event within the event category will be summarized descriptively, for each event occurrence, for patients in the

Double-Blind Safety Population. In addition, patient-level duration of events will be displayed graphically using the MedDRA Preferred Term, spanning the double-blind period. Only patients who have had an adverse event in the given adverse event category will be presented, and multiple occurrences of a single Preferred Term will be presented on the same line.

By-patient listings of deaths, TESAEs, AEs leading to treatment discontinuation, AEOIs, as well as all AEs will be provided. The by-patient AE data listings will include age group, onset and resolution dates, study period, verbatim term, System Organ Class, Preferred Term, period, severity, seriousness criteria, relationship to treatment, action taken for the event, other action, and outcome. The listings will include all data observed across the double-blind and open-label periods, using the Double-Blind Safety Population.

4.11.2 Clinical Laboratory Tests and Electrocardiogram

Clinical laboratory and ECG data will be presented for each treatment group using descriptive statistics based on the observed values and change from baseline values at each scheduled visit of the double-blind period, using the Double-Blind Safety Population. Summary statistics (n, mean, median, SD, minimum, and maximum) will be calculated for quantitative safety data and frequency counts will be compiled for classification of qualitative safety data. This will be repeated for visits corresponding to the open-label period, using the Open-Label Safety Population.

The number and percentage of patients with laboratory values that shift categories (i.e., below/within/above normal range) will be tabulated for each study visit and worst post-baseline visit during the double-blind period. This will be repeated using the Open-Label Safety Population as well as the Fosmetpantotenate Safety Population, for their respective treatment periods.

In addition, the number and percentage of patients with "potentially clinically significant laboratory values" (based on ranges defined below) will be summarized by treatment group, using the Double-Blind Safety Population and the Open-Label Safety Population.

Category	**Parameter**	**Potentially Clinically significant ranges**
Hematology	Hemoglobin	≤ 95 g/L (female), ≤ 115 g/L (male)
	Hematocrit	≤ 0.32 L/L(female); ≤0.37 L/L (male)
	Leukocytes	≤ 2.8 x 10^9/L
	Platelets	≤ 75 x 10^9/L; ≥ 700 x 10^9/L
Coagulation	PT-INR	≥ 2.0
	aPTT	≥ 43 seconds

Category	Parameter	Potentially Clinically significant ranges
Blood Chemistry	Calcium	< 1.75 mmol/L > 3.0 mmol/L
	Chloride	< 90 mmol/L > 115 mmol/L
	Creatinine	> 176.8 umol/L
	Glucose	< 2.8 mmol/L >13.9 mmol/L
	Phosphate	< 0.6 mmol/L > 1.6 mmol/L
	Potassium	< 3.0 mmol/L > 5.5 mmol/L
	Sodium	< 130 mmol/L > 150 mmol/L
	Blood urea nitrogen	> 10.7 mmol/L
Liver Function Tests	Albumin	<25 g/L
	Total Protein	< 45 g/L
	ALT	> 3x-, 5x-, 10x-, and 20xULN
	AST	> 3x-, 5x-, 10x-, and 20xULN
	ALT and AST	Both ALT and AST > 3x-, 5x-, 10x-, and 20xULN
	Total Bilirubin	>ULN and >2xULN
	ALP	>2.5xULN
	ALT and Total Bilirubin	ALT >3xULN + total bilirubin > >2xULN, or international normalized ratio (INR) >1.5
	AST and Total Bilirubin	AST >3xULN + total bilirubin > >2xULN, or international normalized ratio (INR) > 1.5

For ECG measurements, QTc intervals will be determined using the Fridericia correction (QTcF) and the Bazett correction (QTcB). The QTcF and QTcB values and their change from baseline at each scheduled visit of the double-blind period will be summarized using descriptive statistics. The number and percentage of patient with elevated values (per ICH E14 criteria) will also be tabulated. This will be repeated using the Open-Label Safety Population, for visits in the open-label period.

All clinical laboratory safety data, including measurements obtained from the 12-lead electrocardiogram (ECG), will be listed for individual patients across all visits in the double-blind and open-label periods. Values identified as potentially clinically significant (laboratory values) or elevated (ECG values) will be flagged.

4.11.3 Physical Examination and Vital Signs

Reported abnormalities during the double-blind period and open label period, will be tabulated by body system for patients in the Double-Blind and Open-Label Safety Populations, respectively. Changes in physical examinations will be described in the text of the final study report. A by-patient listing will be provided, which presents physical examination results across all assessments in the double-blind and open-label periods.

Vital signs (systolic and diastolic blood pressure, heart rate, respiration rate, and temperature), and body weight will be summarized by treatment group using descriptive statistics. Summary statistics (n, mean, median, SD, minimum, and maximum) will be calculated for both the actual value and the change from baseline value at each scheduled visit during the double-blind period. This will be repeated using the Open-Label safety population for visits during the open-label period.

For systolic and diastolic blood pressure, as well as weight, figures will be presented which show observed values, as well as change from baseline, by treatment group, for each visit during the double-blind period and open-label period, for patients in the Double-Blind and Open-Label Safety Populations, respectively.

In addition, the number and percentage of patients who meet the following criteria at any time during the double-blind period and open-label period will be summarized for each of the respective safety populations:

- Severe blood pressure elevation
    - Systolic blood pressure ≥ 180 mm Hg; **OR**, diastolic blood pressure ≥ 120 mm Hg
- Moderate blood pressure elevation
    - Systolic blood pressure ≥ 140 mm Hg **OR**, diastolic blood pressure ≥ 90 mm Hg

Patients with potentially clinically significant changes from baseline (i.e., elevated values per criteria above) will be summarized and flagged in data listings.

4.11.4 C-SSRS Scale

Columbia Suicide Severity Rating Scale is a suicide assessment questionnaire, which evaluates suicidal ideation and behavior. The questions after the screening visit are all based on “since last visit”.

Descriptive statistics will include the following tabulations by treatment group, with data pooled for pediatric and adult populations, across all visits of the double-blind period:

- The number (%) of patients with suicidal ideation or suicidal behavior at each visit: A "yes" answer at any time during treatment period to any one of the suicidal ideation and behavior questions.
- Shift tables that display the frequency of change in C-SSRS categories during the treatment period relative to baseline at each visit, as well as worst post-baseline visit.
- Severity, frequency, duration, controllability, deterrents, reason (for pediatric patients, only severity and frequency are collected) of suicidal ideation at each visit
- Actual and Potential Lethality/Medical Damage of Attempted Suicide (for actual attempts only) at each visit

The above tables will also be produced for visits in the open-label period, using the Open-Label Safety Population.

Data listings will provide individual patients' responses to the C-SSRS questionnaire at all visits for each patient evaluated on this scale.

4.12 Open-Label Extension

Analyses of data collected during the open-label extension phase will be presented using methodologies similar to those described for the double-blind period. Since the main objective of the open-label extension phase is to evaluate long-term safety, efficacy measurements will mainly be summarized descriptively, with some hypothesis testing for exploratory purposes, as part of the Final Analysis. For the Interim Analysis, open-label safety data will be summarized and presented in listings, while open-label efficacy data will only be presented in listings.

4.12.1 Open-Label Efficacy Analysis

For the Interim Analysis, the scope of analysis of open-label efficacy data is limited to double-blind data. Open-label efficacy data will be included in the listings only.

A later version of the SAP will be released, which will describe the scope of open-label efficacy analysis for the Final Analysis. The analysis is outlined below:

The effects of fosmetpantotenate during the open-label use will be examined based on data from the OL-FAS.

For each efficacy endpoint collected in the open-label extension phase, scores will be tabulated using descriptive summary statistics (i.e., mean, SD, median minimum, and maximum for score totals; frequency tables for ordered categorical ratings) and graphical plots of mean values and mean change from baseline values at each study visit, overall

and by prior treatment in the double-blind period. For this purpose, baseline is defined as the last assessment prior to the first dose of open-label fosmetpantotenate.

On an exploratory basis, hypothesis testing will be applied to the primary and secondary efficacy endpoints measured during the open-label period to assess changes from Baseline, both overall and in the groups to which patients were initially randomized for the double-blind period (i.e., fosmetpantotenate or placebo). Between-group t tests will be used to explore whether improvement is greater for patients who convert to RE 024 from placebo, compared to those maintained on active investigational product for the open-label period. If the number of patients permits, mixed-effect linear models (as described for the primary efficacy analysis) may also be used to explore differences in the pattern of change over time between groups.

An additional exploratory analysis of PKAN-ADL may be undertaken using a mixed model, with Weeks since start of open-label treatment (continuous variable) as the fixed effect, baseline value (for the open-label extension period) as a covariate, and patient as the random effect. The first-order autoregressive AR(1) option will be used for modeling the covariance structure within-patient. The estimate of the time effect (i.e., rate of change) and its 95% CI will be obtained.

For all other efficacy measurements collected in the open-label extension phase, descriptive summary statistics will be presented by study visit. Graphs will be provided for each efficacy endpoint to display mean values (with standard errors) for all patients evaluated at each visit during the open-label extension period.

4.12.2 Open-Label Safety Analysis

For the Interim Analysis as well as the Final Analysis, all safety data collected for the Safety Population during the open-label period of the study will be presented using descriptive statistics, including summary tables of AEs, summary tables and graphical plots of mean values over time for continuous measurements (e.g., vital signs, weight), and frequency tables for categorical data obtained at each study visit. Statistical methods and counting rules similar to those described for the double-blind treatment period will be applied. The last measurement obtained prior to the first dose of open-label fosmetpantotenate will serve as the baseline value for assessing change from baseline in clinical laboratory data and vital signs during the open-label period of the study.

4.13 Pharmacokinetic Analysis

Multiple-dose blood concentrations of fosmetpantotenate will be tabulated and summarized using descriptive statistics. For each analyte, concentration by patient, by nominal sample time, and by study day will be tabulated; summary statistics (n, arithmetic mean, geometric mean, coefficient of variation [CV, as percent], minimum, median, maximum) will be provided by nominal sample time and study day.

Additionally, summary statistics by nominal sample time irrespective of study day will also be provided. Graphical presentations of drug concentration will include individual concentration versus time from dose, by study day and irrespective of study day, and mean concentration versus time from dose, by study day and irrespective of study day. Each graph will be provided on linear and log-linear axes.

Fosmetpantotenate concentrations may also be analyzed by nonlinear mixed-effects modeling, either alone or in conjunction with data from other studies. Available metabolite concentrations may be analyzed similarly.

4.14 Biomarkers

Biomarker assessments (succinate, lactate, maleate, pyruvate, alanine, acetone, acetoacetic acid, 3-hydroxybutyric acid, 2-hydroxybutyric acid, and acetylCoA) will be listed and summarized by treatment group and time point using descriptive statistics (mean, median, standard deviation, standard error of the mean, coefficient of variation, minimum, and maximum) and graphical displays. Change from baseline values will be evaluated at each time point, using paired t-tests (within-group) and 2-sample t tests to evaluate differences between treatment groups. Additional exploratory analyses of biomarkers will be described in a separate SAP.

5 References

1. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH Harmonised Tripartite Guideline, Statistical Principles for Clinical Trials (E9), 5 February 1998.
2. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, ICH Harmonised Tripartite Guideline, Structure and Content of Clinical Study Reports (E3), 30 November 1995.
3. Prakash A., et. al. The impact of analytic method on interpretation of outcomes in longitudinal clinical trials. International J. of Clinical Practice. 2008; 62, 8: 1147-1158.
4. Ouyang J, Carroll KJ, Koch G, Li J. Coping with missing data in phase III pivotal registration trials: Tolvaptan in subjects with kidney disease, a case study. Pharmaceutical Statistics. 2017;16 (4): 250-266
5. Rubin, D.B. (1987) Multiple Imputation for Nonresponse in Surveys. John Wiley & Sons Inc., New York.

6. Mantel, N. and Haenszel, W. Statistical Aspects of Analysis of Data from Retrospective Studies of Disease. Journal of the National Cancer Institute. 1959; 22, 719–748.
7. The EuroQol Group EuroQol-a new facility for the measurement of health-related quality of life. Health Policy 1990; 16(3):199-208.
8. Fischer JS, Jak AJ, Kniker JE, Rudick RA, Cutter G, Multiple Sclerosis Functional Composite (MSFC): Administration and Scoring Manual. 2001 (41 pages)
9. Coleman C, Sobieraj D, Marinucci L. Minimally important clinical difference of the Timed 25-Foot Walk Test: results from a randomized controlled trial in patients with multiple sclerosis. Curr Med Res Opin 2012;28:49–56
10. Westfall PH, Krishen A. Optimal weighted, fixed sequence, and gatekeeping multiple testing procedures. J Stat Planning Inference; 2001; 99:25-40.
11. Mosteller, RD. Simplified calculation of body-surface area. N Engl J Med 1987. 317 (17): 1098.
12. Janssen B, Szende A and Ramos-Goni JM. Self-Reported Population Health: An International Perspective based on EQ-5D. Dordrecht: Springer Open 2014.
13. Choi, S. W., Podrabsky, T., McKinney, N., Schalet, B. D., Cook, K. F., & Cella, D. (2012). PROSetta Stone® Analysis Report: a Rosetta Stone for Patient Reported Outcomes. Volume 1. Chicago, IL: Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University.
14. Brookmeyer R, Crowley JJ. A confidence interval for the median survival time. Biometrics. 1982; 38: 29-41
15. American College of Cardiology (2017, Nov 13) New ACC/AHA High Blood Pressure Guidelines Lower Definition of Hypertension. Retrieved from http://www.acc.org

6 Programming Considerations

All tables, data listings, figures (TLFs), and statistical analyses will be generated using SAS® Version 9.3 (or higher), except for the population pharmacokinetic outputs. Computer-generated output will adhere to the following specifications.

6.1 Table, Listing, and Figure Format

6.1.1 General

1. All TLFs will be produced in landscape format.
2. All TLFs will be produced using the Times New Roman font, size 8 (minimum).

3. The data displays for all TLFs will have a 1.5-inch binding margin on top of a landscape oriented page and a minimum 1-inch margin on the other 3 sides.
4. Headers and footers for figures will be in Times New Roman font, size 8 (minimum).
5. Legends will be used for all figures with more than 1 variable, group, or item displayed.
6. TLFs will be in black and white (no color).
7. Only standard keyboard characters will be used in the TLFs. Special characters, such as non-printable control characters, printer-specific, or font-specific characters, will not be used. Hexadecimal-derived characters will be used, where possible, if they are appropriate to help display math symbols (e.g., μ). Certain superscripts (e.g., cm^2) will be employed on a case-by-case basis.
8. Mixed case will be used for all titles, footnotes, column headers, and programmer-supplied formats.

6.1.2 Headers

All output will have the following header at the top of the page:

Retrophin, Inc. Page n of N
FOSMETPANTOTENATE (Phosphopantothenate Replacement Therapy)
Clinical Study Number 024PKAN15004

All output will have page numbers. TLFs should be internally paginated in relation to the total length (i.e., the page number should appear sequentially as page n of N, where N is the total number of pages in the table).

6.1.3 Display Titles

Each TLF will be identified by a numeral, and the designation (i.e., Table 1) should be centered above the title. A decimal system (14.x-y.z, 14.x.y-z, and 16.2.x-y) will be used to identify TLFs with related contents. The title will be centered in initial capital characters. The analysis set will be identified on the line immediately following the title. The title and table designation will be single spaced. A solid line spanning the margins will separate the titles from the column headers. There will be 1 blank line between the last title line and the solid line.

Table 14.x.y-z
First Line of Title
Second Line of Title if Needed

Full Analysis Set

6.1.4 Column Headers

1. Column headings will be displayed immediately below the solid line described above, in initial upper-case characters.
2. For numeric variables, units will be included in column or row heading when appropriate.
3. Analysis set sizes will be presented for each treatment group in the column heading as (N=xx) (or in the row headings if applicable). This is distinct from the 'n' used for the descriptive statistics representing the number of patients in the analysis set.
4. The order of treatment labels in tables and listings will be: placebo, fosmetpantotenate for the double-blind treatment period; and Open-Label fosmetpantotenate (after Double-Blind placebo), Open-Label fosmetpantotenate (after Double-Blind fosmetpantotenate) for the open-label extension phase; and Total (if applicable for either phase of the study).

6.1.5 Body of the Data Display

1. Listings will be sorted for presentation in order of study phase, treatment group, as above, patient ID, collection day, and collection time.
2. For ordered categorical parameters, frequency tables will present all categories of the parameter between the minimum and maximum level, even if n=0 for all treatment groups in a given category. For example, the frequency distribution for symptom severity would appear as:

Severity Rating	n
severe	0
moderate	8
mild	3

   Where percentages are presented in these tables, any counts of 0 will be presented as 0 and not as 0 (0%).

3. If the categories for a table are not ordered (e.g., Medical History, Reasons for Discontinuation from the Study, etc.), then only those categories for which there is at least 1 patient represented in 1 or more groups will be included.
4. An Unknown or Missing category will be added to the summary table of any categorical parameter for which information is not available for 1 or more patients.
5. Unless otherwise specified, the estimated mean and median for a set of values will be displayed to 1 more significant digit greater than the original values; standard deviation (SD) will be presented to 2 more significant digits than the original values. The minimum and maximum will be reported at the same number of significant digits as the original values. LSmeans will be reported to 1 decimal place, and corresponding CI's to 2 decimal places. For example, for systolic blood pressure:

n	XX
Mean	XXX.X
SD	X.XX
Median	XXX.X
Minimum	XXX
Maximum	XXX

6. Data in columns of a table will be formatted as follows:
   - alphanumeric values left-justified;
   - whole numbers (e.g., counts) right-justified; and
   - numbers containing fractional portions decimal-aligned.
7. Percentage values will be formatted with 1 digit to the right of the decimal point in parentheses, 1 space after the count (e.g., 7 (12.8%), 13 (5.4%)). Less-than signs (e.g., "<0.1%") will appear when values are >0.0% and <0.1% (but not equal to 0.0%). Unless otherwise noted, the denominator for calculating percentages will be the number of patients in a given treatment group and analysis set.
8. Tabular displays of data for prior/concomitant medications and all tabular displays of adverse event data will be presented by body system, drug class, or SOC with the highest occurrence in the fosmetpantotenate treatment group for the double-blind period, or in the group maintained on fosmetpantotenate in the open-label extension phase, in decreasing order. Within the body system, drug class and SOC, medical history (by PT), drugs (by ATC2 code), and adverse events (by PT) will be displayed in decreasing order. If the incidences of multiple terms are identical, they will be sorted alphabetically.
9. Missing data will be left as blank in patient listings.
10. Dates will be formatted in SAS Date9 date format ddmmmyyy (e.g., "27Jan2018"). Missing portions of dates will be represented on patient listings as blank (e.g., "Jan2018"). Dates that are missing because they are not applicable for the particular patient will be presented as "N/A", unless otherwise specified.
11. All observed time values will be presented using a 24-hour clock in hh:mm:ss format (e.g., "01:35:45", "21:26"). Time values will be reported only if they were measured as part of the study.

6.1.6 Footnotes

1. A solid line spanning the margins will separate the body of the data display from the footnotes.
2. All footnotes will be left justified with single-line spacing immediately below the solid line underneath the data display.
3. Each new footnote will start on a new line.
4. Footnotes will appear on each page. Patient specific footnotes will generally be avoided.
5. Footnotes will be used sparingly, and only if they add value to the table, figure, or data listing. If a data display has more than 4 footnotes, they will all appear on a

cover page for the data display; only those footnotes essential for clarification of the data will be repeated on each page.

6. All abbreviations used in a given display will be defined in the corresponding footnote.
7. The last line of the footnote section will be a standard source line, indicating the data source used by the SAS program that produced the data display, the name of the SAS program, and the listing source (e.g., "Data source: xyzabc.sas7bdat Program source: myprogram.sas Listing source: 16.x.y.z").

6.2 Data-Handling Rules

This section describes naming conventions and rules for calculations common to all applicable tables. Some rules specific to a table can be found in the relevant mock-ups.

6.2.1 Unit Conversion to Months/Years

If months are calculated for a duration, then the following conversion is used:

Duration_in_months = Duration_in_days / 30.42

If years are calculated for a duration, then the following conversion is used:

Duration_in_years = Duration_in_days / 365.25

6.2.2 Visits

Relative Study Day: The date of first dose of study treatment is Day 1. A minus (–) sign indicates days prior to date of first dose (e.g., Day -5 represents 5 days before first dose.) There is no Day 0.

If a visit is before the date of first dose, the relative study day is calculated as (Visit Date – Date of First Dose). Otherwise, the relative study day for a specific visit is calculated as (Visit Date – Date of First Dose +1).

- Baseline: For all study variables, baseline for the double-blind period is defined as the last measurement obtained prior to the first dose of double-blind study medication, unless otherwise specified; for the open-label extension study, baseline is the last measurement obtained prior to the first dose of open-label study medication (fosmetpantotenate).
- Post baseline: post baseline visits will be nominal visits, as scheduled in the protocol.

6.2.3 Demographics and Baseline Characteristics

- Height (in cm) = height (in inches) * 2.54
- Weight (in kg) = weight (in lbs) * 0.4535
- Body Mass Index (BMI) (kg/m^2) = weight (in kg) / [height (in meters)] **2

- Body Surface Area[8] (BSA) (m^2) = [height (cm) * weight (kg)/3600]½

6.2.4 Prior and Concomitant Medications

Prior and concomitant medications will be coded and classified using the World Health Organization (WHO) Drug Dictionary Sep 2018 B3 Format.

When partial dates are present in the data, both a partial start date and/or a partial stop date will be evaluated to determine whether it can be conclusively established that the medication either ended prior to first dose of blinded investigational product or started after. If the above cannot be conclusively established based on the partial and/or present dates, then the medication will be counted as concomitant for the double-blind period and open-label period (if the patient has reached the open-label period).

The imputation method for the handling of missing or partial dates will be as follows. When the start date and the stop date are both incomplete for a patient, the start date will be imputed first:

Partial start date

Missing Month and Day

- If the year of the partial start date is the same as the year of the first dose of blinded investigational product, then the month and day of the first dose of blinded investigational product will be assigned to the missing fields.
- If the year of the partial start date is before the year of the first dose of blinded investigational product, *December 31* will be assigned to the missing fields.
- If the year of the partial start date is after the year of the first dose of blinded investigational product, *January 1* will be assigned to the missing fields.

Missing Month Only

- If only the month is missing, then the month only will be imputed according to the procedure above.

Missing Day Only

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

blinded investigational product, then the first day of the month will be assigned to the missing day.

If the stop date is complete and the imputed start date as above is after the stop date, then the start date will be imputed with the stop date.

Partial stop date

Missing Month and Day

- If the year of the partial stop date is the same as the year of the last dose of investigational product (double-blind or open-label), then the month and day of the last dose of investigational product (double-blind or open-label) will be assigned to the missing fields.
- If the year of the incomplete stop date is before the year of the last dose of investigational product (double-blind or open-label), *December 31* will be assigned to the missing fields.
- If the year of the incomplete stop date is after the year of the last dose of investigational product (double-blind or open-label), *January 1* will be assigned to the missing fields.

Missing Month Only

- If only the month is missing, then the month only will be imputed according to the above procedure.

Missing Day Only

- If the month and year of the partial stop date are the same as the month and year of the last dose of investigational product (double-blind or open-label), then the day of the last dose of investigational product (double-blind or open-label) will be assigned to the missing day.
- If either the year of the incomplete stop date is before the year of the date of the last dose of investigational product (double-blind or open-label) or if both years are the same but the month of the partial stop date is before the month of the date of the last dose of investigational product (double-blind or open-label), then the last day of the month will be assigned to the missing day.
- If either the year of the partial stop date is after the year of the date of the last dose of investigational product (double-blind or open-label) or if both years are the same but the month of the partial stop date is after the month of the date of the last dose of investigational product (double-blind or open-label), then the first day of the month will be assigned to the missing day.

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.

6.2.5 Adverse Events

Adverse events will be coded and classified using the MedDRA coding dictionary Version 20.0. The specific dictionary version will appear in the actual tables/listings.

AEs with missing start and end dates will be considered as treatment emergent for the open-label period, taking a "worst-case" approach. If the patient has not yet reached the open-label period, or if the AE has a stop date overlapping with the double-blind study period, then the AE will be counted as treatment-emergent for the double-blind period.

PAREXEL International Electronic Signature Page

This page is the manifestation of the electronic signature(s) used in compliance with PAREXEL International's electronic signature policies and procedures and in compliance with applicable regulations.

UserName: PPD
Title: PPD
Date: Monday, 22 April 2019, 05:21 PM GMT Standard Time
Meaning: Document contents authored.
==========

UserName: PPD
Title: PPD
Date: Wednesday, 24 April 2019, 01:41 AM GMT Standard Time
Meaning: Document contents reviewed.
==========

UserName: PPD
Title: PPD
Date: Monday, 29 April 2019, 11:18 AM GMT Standard Time
Meaning: Document contents reviewed.
==========

UserName: PPD
Title: PPD
Date: Monday, 29 April 2019, 01:22 PM GMT Standard Time
Meaning: Document contents reviewed.
==========

UserName: PPD
Title: PPD
Date: Monday, 29 April 2019, 02:09 PM GMT Standard Time
Meaning: Document contents approved.
==========