A randomised, double-blind, placebo-controlled study to evaluate the efficacy and safety of safinamide, as add-on therapy, in idiopathic Chinese Parkinson's Disease (PD) patients with motor fluctuations treated with stable doses of levodopa

ClinicalTrials.gov Identifier: NCT03881371

Statistical Analysis Plan: Final 2.0, 08 Nov 2021

PAREXEL International

Zambon S.p.A

Z7219L05

A Randomised, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of Safinamide, as Add-On Therapy, in Idiopathic Chinese Parkinson's Disease (PD) Patients with Motor Fluctuations Treated with Stable Doses of Levodopa

Statistical Analysis Plan

Version: 2.0

PAREXEL Project Number: CCI

SPONSOR SIGNATURE PAGE

Approved by: PPD

PPD

Zambon S.p.A.

PPD

Date

PAREXEL SIGNATURE PAGE

Signature(s) below confirm that the Statistical Analysis Plan was developed in accordance with CCI and that it is approved for release.

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

	Signatory
Author	PPD Project Role: PPD

TABLE OF CONTENTS

REVISION HISTORY

Version No.	Effective Date	Summary of Change(s)
0.1	20 Nov 2018	New document
0.2	26 Nov 2018	Minor updates based on sponsor's comments (only for CTA submission).
0.3	24 Feb 2020	Revised from draft version for CTA submission (draft 0.2), based on final protocol 2.0, 19 Dec 2018.
0.4	06 Mar 2020	Revised based on internal review.
0.5	19 Mar 2020	Updated Sponsor comments.
1.0	30 Mar 2020	Finalized document.
1.1	23 July 2020	• Update reference to CRF version from version 3.0 to 4.0. • In section 4.2, revised the definition of study day considering dosing date (i.e. Assessment Date – First Dosing Date + 1) instead of considering randomization date. • In section 4.8, added clarifications on the definition of the total amount of exposure doses (mg) and number of scheduled doses. Also added treatment compliance adjustment considering COVID-19 impact. • In section 4.9.1.2, added clarifications on the management of partially or completely missing data on diaries. • In section 4.10.1, clarified definition of TEAE to include those AE occurring after last dose of IMP.
1.2	21 Dec 2020	• In section 4.9.1.2, added another sensitivity to evaluate general impact caused by COVID-19.
1.3	28 Jan 2021	• In section 4.5.5, clarified PP population to include ITT subjects who have to complete the double blind 16-week treatment period and had no major protocol deviations that were considered as potentially impacting the efficacy results. • In section 4.9.1.1, clarified how results should be considered in relation to efficacy results when treatment and centre interaction is statistically significant. • In section 4.9.1.2, clarified FAS Set analysis (ANCOVA, Multiple Imputation) is the primary analysis result according to protocol; Missing data for subjects in PP Set will be imputed by Multiple Imputation. • For all efficacy endpoints, clarifying analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) are considered as exploratory results.

1.4	27 Apr 2021	• In section 4.9.1.2, clarified how to deal with Proc MIANALYZE option when none of subject's Diary or NRS data is missing but need to analyze under PP Set (ANCOVA, Multiple Imputation). • In section 4.9.1.2, clarified MCMC procedure will be applied to proc MI under multiple imputation. • In section 4.9.3.2, clarify the primary conclusion of CGI-C and CGI-S summary will be based on stratified Wilcoxon-Mann-Whitney test. • In section 4.12, the change from protocol planned analysis is updated.
2.0	Date of last signature	

LIST OF ABBREVIATIONS

Abbreviation / Acronym	Definition / Expansion
ADL	Activities of daily life
AE	Adverse event
ANCOVA	Analysis of covariance
ATC	Anatomical Therapeutic Chemical
BLQ	Below the lower limit of quantification
BMI	Body Mass Index
BP	Blood pressure
CGI-C	Clinical Global Impression of Change
CGI-S	Clinical Global Impression of Severity
CI	Confidence interval
CGI	Clinical Global Impression
CRF	Case Report Form
CRO	Contract research organization
CSP	Clinical Study Protocol
CV	Coefficient of variation
DBL	Database Lock
DNA	Deoxyribonucleic acid
DOV	Date of Visit
DRM	Data Review Meeting
ECG	Electrocardiogram
EMA	European Medicines Agency
ET	Early termination
FDA	Food and Drug Administration
ICF	Informed consent form
IMP	Investigational Medicinal Product
MedDRA	Medical Dictionary for Regulatory Activities
NA	Not available
NK	Not known

Abbreviation / Acronym	Definition / Expansion
NRS	Numerical Rating Scale
PD	Parkinson's Disease
PE	Physical Examination
PDQ-39	Parkinson's Disease Questionnaire-39 items
PP	Per-Protocol Population
PT	Preferred Term
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SE	standard error
SD	Standard Deviation
SOC	System Organ Class
TEAE	Treatment-emergent adverse event
UPDRS	Unified Parkinson's Disease Rating Scale
WHODrug Global	World Health Organization - Drug Global

1 INTRODUCTION

Safinamide has been approved by the European Medicines Agency (EMA) for the treatment of mid- to late-stage fluctuating PD patients as add-on therapy to L-dopa (alone or in combination with other anti-Parkinson drugs) and by the Food and Drug Administration (FDA) as adjunctive treatment to levodopa/carbidopa in patients with Parkinson's Disease experiencing "OFF" episodes. At the current date safinamide is on the market in 11 European countries, in Switzerland and in US.

The efficacy of safinamide has been demonstrated in patients with motor fluctuations when administered as add-on therapy alongside standard of care therapy including L-dopa, dopamine agonists, catechol O-methyltransferase inhibitors, anticholinergics and amantadine, thus emphasizing the additional benefits it can offer when patients are no longer optimally controlled on their current treatment regimen. Importantly, a notable improvement in motor fluctuations is achieved without an increase in troublesome dyskinesia and the benefits are long lasting.

This study is designed to collect data on the impact of safinamide on the motor complications of Chinese PD fluctuating patients over a treatment period of up to 16 weeks. Further information on the effect of safinamide on motor symptoms and quality of life in a clinical setting environment will also be gathered.

The analyses described in this SAP are based upon the following study documents:

- Amended Study Protocol, Version 2.0 (October 08, 2019)
- Study Case Report Form, Version 4.0 (April 27, 2020)

The Statistical Analysis Plan (SAP) details the statistical methodology to be used in analyzing study data and outlines the statistical programming specifications for the Tables, Listings, and Figures. It describes the variables and populations, anticipated data transformations and manipulations and other details of the analyses not provided in the Clinical Study Protocol (CSP).

The SAP will be finalized prior to database lock (DBL) and describes the statistical analysis as it is foreseen when the study is being planned. If circumstances should arise during the study rendering this analysis inappropriate, or if improved methods of analysis should arise, updates to the analyses may be made. Any deviations from the SAP after DBL, reasons for such deviations and all alternative or additional statistical analyses that may be performed, will be described in a SAP Addendum.

2 STUDY OBJECTIVES

2.1 Efficacy Objective

The objective of the study is to evaluate the efficacy of safinamide compared with placebo, given as add-on therapy, in idiopathic Chinese PD patients with motor fluctuations treated with stable doses of levodopa (L-dopa).

2.2 Safety Objective

The safety objective of the study is to evaluate the safety and tolerability of safinamide compared with placebo in Chinese PD patients with motor fluctuations.

3 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

This is a Phase III, multicentre, randomised, double-blind, placebo controlled study in idiopathic Chinese PD patients, experiencing motor fluctuations while on stable doses of levodopa (alone or in combination with other anti-Parkinson drugs).

A total of 306 patients will be randomised into this study (153 in the safinamide and 153 in the placebo groups).

The visit procedures are summarised in the Study Flow Chart, Appendix 1.

Screening Period

After providing written informed consent to participate in the study, patients will enter a screening period up to 2 weeks. During the screening period, patients will undergo all the evaluations necessary to establish their eligibility for the study. Patients considered non-eligible ("screening failures") due to clinically significant abnormalities in laboratory exams, ECG, vital signs or traditional Chinese medicine not related to nervous system disease, could be re-screened again only once during the study after a reasonable interval of time, based on the judgement of the Investigator that should be fully documented and explained in the clinical records and in the CRF. They would need to sign a new consent form. Patients who will be confirmed to be non-eligible in this second screening visit should be definitively excluded from the study.

Patients and their caregivers will be trained on the completion of a 24-hour diary card and the last two days of recording prior to each study visit will be used for data analysis. The dose of L-dopa and of the other anti-Parkinson drugs (if any) must be kept constant during the screening period.

Treatment Period

At baseline (day 1), eligible patients will enter the treatment period and will be randomised to receive either safinamide (initial 50 mg titrated to 100 mg the day after the Visit 3/week 2, ideally at day 15) or matching placebo, orally od in a 1:1 ratio. The investigational medicinal product (IMP)

will be taken in the morning at breakfast time, in addition to the morning dose of L-dopa and other (if any) PD medications.

Following completion of all baseline assessments, they will receive the first dose of safinamide or placebo (50 mg) at the study centre. The day after the Visit 3/week 2 (ideally at day 15) the dose will be increased at home to 100 mg od. Each patient will receive treatment for 16 weeks, with visits at week 0/day 1 (baseline) and at weeks 2, 6, 10 and 16 (or early termination). A telephone follow-up will be performed 1 week after the end of treatment for safety reasons.

Patients who prematurely withdraw from the study while receiving study medication should complete the early termination visit assessments, when possible.

At the end of the study, the patients will be instructed to contact immediately the Investigator in case of appearance of any adverse reactions. Any ongoing adverse event or clinically abnormal laboratory parameter will be followed until resolution. In addition, all SAEs occurring within 30 days after a patient's last dose of study drug will be followed to their conclusion.

In the case of intolerable dopaminergic adverse events (AEs), e.g. dyskinesia, it is suggested to decrease the dose of L-dopa by a telephone call as first step and consider the decrease of the dose of safinamide or placebo from 100 to 50 mg od as a second step. In this second case, patients should undergo an unscheduled visit for safety reasons and maintain the 50 mg safinamide dose for the rest of the study. Patients who do not tolerate the 50 mg dose should withdraw from the study and complete the early termination visit assessments, when possible.

Efficacy will be assessed by the changes in "OFF" and "ON" time from the 24-hour patient diary, the Unified Parkinson's Disease Rating Scale (UPDRS), the Clinical Global Impression (CGI), the Parkinson's Disease Questionnaire-39 items (PDQ-39) and the Numerical Rating Scale (NRS).

Safety will be assessed by clinical laboratory tests (haematology and serum chemistry), vital signs, 12-lead electrocardiogram, physical examination, treatment emergent adverse events and concomitant medications.

3.2 Endpoints

3.2.1 Efficacy Variables

3.2.1.1 Primary Endpoint

The primary efficacy endpoint of this study is the change from baseline to week 16 in the mean total daily "OFF" time, as assessed by 24-hour patient diary cards, of safinamide 100 mg/day compared to placebo, given as add-on therapy in PD patients with motor fluctuations on stable doses of L-dopa.

3.2.1.2 Secondary Endpoints

The key secondary efficacy endpoint of this study is:

- The change from baseline to week 16 in pain severity, as assessed by an 11-point Numerical Rating Scale (NRS).

Other secondary efficacy endpoints of this study are:

- The change from baseline to week 16 in the mean total daily "ON" time, as assessed by 24-hour patient diary cards.
- The change from baseline to week 16 in the mean daily "ON" time with no/non-troublesome dyskinesia, as assessed by 24-hour patient diary cards.
- The change from baseline to week 16 in the Unified Parkinson's Disease Rating Scale (UPDRS) total score during the "ON" phase.
- The change from baseline to week 16 in the UPDRS part II (activities of daily life [ADL]) score during the "ON" phase.
- The change from baseline to week 16 in the UPDRS part III (motor function) score during the "ON" phase.
- The CGI-Severity (CGI-S) score at week 16.
- The change from baseline to week 16 in the CGI-Change (CGI-C).
- The change from baseline to week 16 in the PDQ-39 score.

3.2.2 Safety Endpoints

The safety endpoints for this study are:

- The nature, frequency, severity, relationship (to study drug), actions taken, and outcome of treatment-emergent adverse events (TEAEs) and treatment-emergent serious adverse events (TSAEs).
- Physical examination findings (clinically significant).
- Vital sign (heart rate, systolic and diastolic blood pressure) values, including occurrence of abnormalities.
- 12-lead electrocardiogram (ECG) parameter measures, including occurrence of abnormalities
- Clinical chemistry and haematology values, including shifts from baseline and occurrence of abnormalities.

4 STATISTICAL METHODS

4.1 Data Quality Assurance

All tables, figures and data listings to be included in the report will be independently checked for consistency, integrity and in accordance with standard PAREXEL procedures.

4.2 General Presentation Considerations

'Baseline' is defined as the last available pre-treatment assessment at week 0/day 1. 'Study Day' will be calculated relative to the date of first dosing i.e. Study Day = Assessment Date - First Dosing Date + 1.

Continuous data will be summarized in terms of the mean, standard deviation (SD), median, minimum, maximum and number of observations, unless otherwise stated. The minimum and maximum will be reported to the same number of decimal places as the raw data recorded in the

database. The mean and median will be reported to one more decimal place than the raw data recorded in the database. The SD will be reported to two more decimal places than the raw data recorded in the database. In general, the maximum number of decimal places reported shall be four for any summary statistic.

Categorical data will be summarized in terms of the number of subjects providing data at the relevant time point (n), frequency counts and percentages. Any planned collapsing of categories will be detailed in the SAP text and the data displays.

Percentages will be presented to two decimal places. Percentages will not be presented for zero counts. Percentages will be calculated using n as the denominator. If sample sizes are small, the data displays will show the percentages, but any textual report will describe frequencies only.

Changes from baseline in categorical data will be summarized using shift tables where appropriate.

P-values greater than or equal to 0.001, in general, will be presented to three decimal places. P-values less than 0.001 will be presented as "<0.001".

Confidence intervals will be presented to one more decimal place than the raw data.

All data captured on eCRFs will be available as listings.

Unless stated otherwise, all available data from withdrawn subjects will be included in the analysis up to the time of withdrawal.

4.3 Software

All report outputs will be produced using SAS® version 9.4 or a later version in a secure and validated environment.

4.4 Study Subjects

4.4.1 Disposition of Subjects

A clear accounting of the disposition of all subjects who enter the study will be provided, from screening to study completion. Subjects are allowed to be re-screened; all screen failure data which are entered in database will be utilized for summary.

The following summaries will be provided for Enrolled Population:

- Subjects screened
- Subjects excluded prior to randomization by major reason

The following summaries will be provided by treatment group and overall for Randomized Population:

- Subjects Randomized
- Subjects Completed the Study
- Subjects Treated (with at least one dose or partial dose of study medication)
- Subjects withdrawn from treatment by major reason and overall
- Subjects withdrawn from the study by major reason and overall

The following by-subject listings will be provided for Enrolled Population:

- Subjects screened with screen failure flag and the major reason of exclusion prior to randomization.

The following by-subject listings will be provided for Randomized Population:

- Randomization code assigned treatment and actual treatment.
- Withdrawal/study completion details (including reason for discontinuation and duration of treatment prior to discontinuation).

4.4.2 Protocol Deviations

Major protocol deviations are defined as those deviations from the protocol likely to have an impact on the perceived efficacy and/or safety of study treatments. The impact of major protocol deviations on the efficacy and/or safety results will be investigated by assessing the robustness of the study results and conclusions to the choice of analysis population, both including and excluding data potentially affected by major protocol deviations. For the definitions of major protocol deviation, please refer to Section Per-Protocol Population.

Major protocol deviations and any action to be taken regarding the exclusion of subjects or affected data from specific analyses are defined in the project-specific Protocol Deviation Specification.

A summary of the number and percentage of subjects with a major or major COVID-19 protocol deviation by treatment group and overall and by type of deviation will be provided for Randomized Population by treatment group and overall.

A by-subject listing of major and minor protocol deviations will be provided.

4.5 Analysis Populations

The following analysis will be provided for Enrolled Population:

- A summary of the number and percentage of subjects by treatment group and overall for each analysis population.
- A by-subject listing of analysis population will include: treatment group, center, subject identifier, inclusion/exclusion flag for each population and reason for exclusion from each population.

4.5.1 Enrolled Population

The Enrolled Population will include all subjects who provide informed consent.

4.5.2 Randomized Population

The Randomized Population will include all subjects who provide informed consent and received a patient number (randomization number) whether or not they receive IMP.

4.5.3 Full Analysis Set (FAS) Population

The FAS population will comprise all patients who provided informed consent, were randomized and received at least 1 dose or partial dose of the IMP.

Primary analyses will be performed on the FAS population with exclusions from the randomized defined and justified in the SAP.

Following the randomized principle, patients will be analyzed according to the treatment they have been assigned to at the randomization.

The FAS population will be used to produce summaries of baseline patient characteristics and for the analysis of all efficacy endpoints.

4.5.4 Safety Population

The Safety Population will comprise all patients who provided Informed Consent and received at least 1 dose or partial dose of IMP.

Patients will be analyzed according to the treatment they actually received.

The Safety Population will be used to produce summaries of all safety related endpoints and demography.

4.5.5 Per-Protocol Population

The Per-protocol Population (PP) will include all FAS population patients who completed the double blind 16-week treatment period and had no major protocol deviations that were considered as potentially impacting the efficacy results.

Major protocol deviations might include, but are not limited to:

- Patients taking a not-permitted concomitant medication.
- The IMP not being administered during the trial as defined in the protocol.
- Patients receiving a treatment different than the one assigned by randomization.
- The treatment compliance less than 80% or greater than 120%.

Results of the primary and secondary efficacy endpoints analyses conducted in the PP will be considered as supportive.

Exclusion of patients from the PP analyses will be decided jointly by the contract research organization (CRO) and Sponsor's Medical Monitor, Clinical Trial Manager and Statistician in the Data Review Meeting (DRM) prior to unblinding of the randomization code and database release.

The patients or observations to be excluded, and the reason for their exclusion will be documented and approved by the above-mentioned persons prior to database release. The documentation will be filed together with the remaining trial documentation.

Violations excluding patients from any particular population will be described, reporting the number of protocol violators per each criterion. All protocol violations, minor ones included, will be listed.

4.6 Demographic and Other Baseline Characteristics

Demographic and other baseline characteristics will be summarized by treatment group and overall for Safety Population. The summaries provided may include the following:

- A summary of demographic variables (age, sex, and ethnicity, baseline height, baseline weight, baseline BMI) by treatment group and overall.
- A summary of Substance Use at Screening (smoking <10 or >=10 cigarettes/Day, alcohol intake: category of "<40g/day, 40-80g/day or >80g/day) by treatment group and overall.
- A summary of Hoehn and Yahr stage at Screening by treatment group and overall.
- A summary of "normal", "abnormal, not clinically significant (NCS)" or "abnormal, clinically significant (CS)" neurological examinations at Screening by each parameter, treatment group and overall.
- A summary of medical history by System Organ Class (SOC) and Preferred Term (PT) by treatment group and overall. The medical history will be coded by MedDRA (version 22.0).
- A summary of surgical history by SOC and PT by treatment group and overall. The surgical history will be coded by MedDRA (version 22.0).

All demographics and other baseline characteristics will be provided in data listing. For Hoehn and Yahr scale and neurological examinations, treatment group actually dosed by subject are provided in data listing.

4.7 Concomitant Medication and Procedure

All medications will be coded using the World Health Organization-Drug Global (WHODrug Global) (version March 2019 B3 format) and will be classified by Anatomical Therapeutic Chemical (ATC) categories.

Prior and concomitant medication start and stop dates will be compared to the date of first dose of study medication to allow medications to be classified as either Prior only or Concomitant only. Medications starting after the study completion/withdrawal date will be listed but will not be classified or summarized.

Medications that stop prior to the date of first dose of study medication will be classified as "Prior" only. Medications will be classified as Concomitant if they have a start date on or after the date of first dose of study medication. In addition, medications stat prior to first dose of study medication and stop after first dosing date are classified as Concomitant medication as well. Imputation of partial dates of medication records for classification of "prior" and "concomitant" is described in Appendix 2.

The concomitant procedures include any surgical, therapeutic or diagnostic procedures performed during study and they will be coded by MedDRA (version 22.0).

The summaries provided may include the following:

- Prior medications will be summarized by ATC level 3 and preferred name by treatment group and overall for Safety Population.
- Concomitant medications will be summarized by ATC level 3 and preferred name by treatment group and overall for Safety Population.
- Concomitant procedures will be summarized by SOC and PT by treatment group and overall for Safety Population.
- By-subject listings of all medications will be provided with flag of "prior" or "concomitant".
- By-subject listings of all concomitant procedures.

4.8 Treatment Compliance

Exposure data will be summarized and listed based on the Safety Population. The compliance data will be summarized and listed based on the Safety Population.

Treatment duration, total amount of exposure doses, compliance to the study drug, and compliance to the study drug by category (<80%, 80%-120%, >120%) will be descriptively summarized by treatment group and overall.

Investigator will assess adherence with IMP dosing regimen on an ongoing basis by determining the amount of IMP dispensed, used and returned at Visit 2 (first dose), Visit 3, Visit 4, Visit 5 and Visit 6 (end of treatment) or unscheduled Visit/premature discontinuation visit.

Treatment duration (weeks) = (date of last dose – date of first dose + 1) /7.

Total amount of exposure doses (mg) = $\sum$ *dose amount actual taken between visits*.

Due to COVID-19, subjects may not be able to return site for scheduled visits as planned and cannot return study drug on schedule visits. Considering this situation, site will contact subjects on planned schedule visit date by phone to record how many tablets of study drug have been taken until date of phone contact. The date of phone contact will be recorded as Date of Visit (DoV) in CRF. This number of tablets taken will be used for summary of total amount of exposure doses (mg).

If subject cannot receive study drug on time due to COVID-19, the remaining study drugs from last visit will be taken until new study drugs are delivered.

Regarding the Returned Date and Number of Tablets Returned in CRF, site will record exactly date subjects go back to site and number of tablets returned.

Compliance to the IMP (%) = $100 \times \frac{total\ amout\ of\ exposure\ doses\ (mg)}{total\ amount\ of\ scheduled\ doses\ (mg)}$.

In protocol, the total amount of scheduled dose is planned as: 50 mg (Week 0 to Week 2) and 100 mg (Week 2 to Week 16). If subject received this dosing schedule of Safinamide per protocol without dose down titration, then the total amount of scheduled dose (mg) in compliance can be calculated as:
50 mg × number of days scheduled between first dosing date and Week 2 + 100 mg × number of days scheduled between Week 2 and End of treatment date.

If subjects need to adjust dose of Safinamide, the total amount of scheduled dose (mg) should be calculated based on adjusted dose.

Considering subjects may be impacted by COVID-19, the total amount of scheduled doses will be calculated based on DOV.

The compliance calculation will be based on visit basis and overall as following:
- Treatment Compliance – Visit 2 to Visit 3 (Week 0 to Week 2)
- Treatment Compliance - Visit 3 to Visit 4 (Week 2 to Week 6)
- Treatment Compliance - Visit 4 to Visit 5 (Week 6 to Week 10)
- Treatment Compliance - Visit 5 to Visit 6 (Week 10 to Week 16)
- Overall Treatment Compliance - Visit 2 to Visit 6 (Week 0 to Week 16)

The following summaries will be provided:
- A summary of the treatment duration and the total amount of exposure doses (mg), and total amount of scheduled doses (mg) administered during study by treatment group and overall.
- A summary of number and percentage of subjects with treatment compliance by treatment group and overall; the treatment compliance by visit is summarized as well.

The following listings will be provided:
- By-subject listing of treatment compliance.
- By-subject listing of drug accountability.

4.9 Efficacy Evaluation

4.9.1 Analysis and Data Conventions

4.9.1.1 Multi-center Studies

For the summaries and analyses, the term ‘Center’ will be used to define each investigator site.

As additional analysis, An analysis of covariance (ANCOVA) models for the analysis of primary endpoint and key secondary endpoint will be fitted again including the site-by-treatment interaction, in order to test if that interaction is statistically significant, using an alpha = 0.1. Also, in order to graphically assess the possible heterogeneity of the treatment effect across centres, forest plots will be generated to display the results at each center. In case that the number of patients in each site were scarce, sites will be gathered into "Region". Upon completion of the study and prior to unblinding, study statistician in consultation with the clinical team will determine the pooling based on enrolment numbers and geographical proximity.

If the site-by-treatment interaction is statistically significant, it implies that the effect of site-by-treatment interaction exists. The overall conclusion of the study on the primary and key secondary endpoints analyses would be still analyzed based on the primary model (i.e., the ANCOVA model without the interaction term). However, an additional exploratory analysis to identify which sites may have significant treatment effects which causes significant interaction is suggested to be shown at the end.

4.9.1.2 Handling of Dropouts or Missing Data

At each Visit, there are 2 diary entries (1 Day or 2 Days Before Visit), and the diary covers 24h (48 entries, each for 0.5h). A missing diary will be classified as follows:

- Complete Missing

If the diary is completely missing, meaning that there are no data collected, on 1 Day or 2 Days Before Visit, then it is defined as "Complete Missing" in that day.

- Partial Missing

If more than 2.5 hours (≥ 5 ticks within 24 entries per day) are missing on 1 Day or 2 Days Before Visit, the diary for that day will be treated as partial missing.

In case of complete missing diaries in both the two days before a visit (both Day 1 and Day 2 Before Visit), or in the case of partial missing diaries in both the two days before a visit (both Day 1 and Day 2 Before Visit), the computation of the endpoint based on diaries data cannot be performed and the diaries are considered missing for the visit.

In case of complete missing or partial missing diaries for a single day before the visit, and availability of diaries data for the other day before visit, the endpoint derivation based on diaries data can be performed using data collected in the available day.

In case of less than 5 missing ticks, the computation of the endpoint will be performed by using the remaining available data.

The same rule to define missing for the visit will be applied to primary and secondary endpoints based on diary data.

Missing data on the primary and key secondary efficacy endpoint will be imputed using multiple imputation (MI) as the primary imputation method and last observation carried forward (LOCF) as sensitivity analysis.
Missing data on all the other secondary efficacy endpoints will be imputed only using LOCF method.

A MI analysis will be carried out using the FAS population. The missing values of mean "OFF" time will be imputed with multiple imputation methodology using Procedure proc MI in SAS use Markov Chain Monte Carlo (MCMC) which assumes that all the variables in the imputation model have a joint multivariate normal distribution. This procedure will be executed by treatment group with a seed of 237006 and will include Baseline mean "OFF" time and each post-baseline visits during the treatment period.

The 50 times of repeated imputed datasets will be generated for this analysis; only on-treatment observations are used. Each of the imputed datasets will be analyzed by timepoint using an ANCOVA model with the same covariates as the original model. The MIANALYZE procedure in SAS will be used to combine the ANCOVA results from 50 imputed datasets at each time point.

Missing data will be imputed using the LOCF method followed by an ANCOVA analysis using the same model as described in primary endpoint. The LOCF should be performed from first post-dose visit until the last scheduled visit (Week 16) to impute the missing values (missing due to treatment discontinuations or intervention censoring, or any other reasons).

Finally, additional sensitivity analysis of general impact by COVID-19 is considered. Based on Protocol Deviation Report, subjects with Major COVID-19 related Protocol Deviation will be excluded from this sensitivity analysis to check if any COVID-19 related Protocol Deviation may bias primary and key secondary efficacy results.

The following efficacy analysis will be provided:

- FAS Set analysis (ANCOVA, Multiple Imputation): this is the primary analysis for primary endpoint (the mean total daily "OFF" time) and key secondary endpoint (NRS score).
- PP Set analysis (ANCOVA, Multiple Imputation): this is the supportive analysis for primary endpoint (the mean total daily "OFF" time) and key secondary endpoint (NRS score).
- FAS Set analysis (ANCOVA, LOCF): this is a sensitivity analysis for all efficacy endpoints.
- FAS Set analysis (ANCOVA, Multiple Imputation, general COVID-19 impact): this is a sensitivity analysis to exclude subjects with Major COVID-19 Protocol Deviation for primary endpoint (the mean total daily "OFF" time) and key secondary endpoint (NRS score).

When there is no missing diary or NRS results under PP set, it will cause warning message and cannot enable SAS Proc MIANALYZE option to return final estimation of least square mean between group and within group across 50 imputations. To overcome this issue, the observed (non-imputed) diary or NRS results in PP set will be used to provide least square means instead of using multiple imputation results.

4.9.1.3 Multiplicity

The overall type I family-wise error rate for testing the primary and the key secondary efficacy endpoints will be controlled at the two-sided 0.05 significance level using a Fixed Sequence Procedure. Following this procedure, progression to next step will only occurs as long as null hypothesis (H_0: Active effect = Vehicle effect) from previous step is rejected at 0.05 significance level. The testing procedure is stopped, and all the remaining null hypotheses accepted if an acceptance occurs. All individual hypothesis tests will be 2 sided.

1. The first step will test the primary efficacy parameter. The p-value for the null hypothesis must be less than 0.05 to be considered to have met the primary efficacy objective. If the null hypothesis is not rejected (i.e., p-value >0.05), the subsequent statistical test (second step) will not be considered statistically significant.

2. The second step will test the key secondary efficacy parameter "change from baseline to week 16 in pain severity, as assessed by an 11-point Numerical Rating Scale".

For the analyses of supportive secondary efficacy endpoints (not key secondary endpoints), **no adjustment of significance level** will be made to account for multiple comparisons.

4.9.1.4 Interim Analyses

There is no interim analysis.

4.9.1.5 Efficacy Data

The FAS population will be used for the primary analyses of each of the efficacy endpoint whilst results from supplemental analyses using the Per Protocol will be compared to those based on the FAS population to assess the effects of dropouts, missing data, protocol violations and deviations.

The distributions of all the efficacy endpoints will be summarized by treatment group and time point. All statistical testing results by visits prior to Visit 6 (Week 16) between treatment groups are provided as supportive results. Counts and percentages will be reported with the latest computed based on the numbers of patients with non-missing observations. The percentages will be suppressed when the count is zero in order to draw attention to the non-zero counts. Furthermore, efficacy endpoints will be further summarized by arithmetic means, standard deviations, median quartiles, minima and maxima.

Under ANCOVA model, the adjusted least squares mean change from baseline by visits will be reported along with the standard error (SE) and 95% confidence interval by each treatment group. For treatment group comparison, the treatment difference between Safinamide versus Placebo will be reported with the adjusted least squares mean difference, SE of difference, 95% confidence interval of the difference along with the p-value. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

Adjusted change from baseline to scheduled visits in primary or key secondary endpoint will be plotted by line plot, showing mean change from baseline with SE.

A by-subject data listing will be provided for all efficacy data.

4.9.2 Primary Efficacy Variable – mean total daily "OFF" time

The diary is completed by the subject two days before the baseline visit (Visit 2), weeks 2, 6, 10 and 16 (visits 3, 4, 5 and 6) for recording of "OFF" and "ON" time for each 30-minute time period. The daily "OFF" time is summation of all "OFF" hours. All "OFF" hours on two days before each scheduled visit will be averaged to get a "OFF" time by visit.

Then mean total daily "OFF" time is averaged over the 2 days prior to each study visit. Taking baseline visit as example:

- 2 Days before Baseline – summation of total "OFF" time in hour and defined as "OFF1-Baseline" .
- 1 Day before Baseline – summation of total "OFF" time in hour and defined as "OFF2-Baseline" .
- Subject's mean total daily "OFF" time at Baseline: average of "OFF1-Baseline" and "OFF2-Baseline".
- Following the above rules to define averaged total daily "OFF" time at Visit 3 (Week 2), 4 (Week 6), 5 (Week 10) and 6 (Week 16) by each subject.

The change from baseline in mean total daily "OFF" Time is calculated by taking the difference between the average of the total daily "OFF" time at Visit 3, 4, 5 and 6, and the baseline total daily "OFF" time.

The primary endpoint is defined as follow:
Change from baseline to week 16 in the mean total daily "OFF" time, as assessed by the 24-hour patient diary, of safinamide 100 mg/day compared to placebo, given as add-on therapy in PD patients with motor fluctuations on stable doses of L-dopa.

The null hypothesis for the primary endpoint comparison will be that there is no difference between safinamide 100 mg/day and placebo, given as add-on therapy in PD patients with motor fluctuations on stable doses of L-dopa, in change from baseline to week 16 in the mean total daily "OFF" time (defined as μ in statistical hypothesis below), as assessed by the 24-hour patient diary. μ is defined as the mean change from baseline to week 16 in the mean total daily "OFF" by treatment group (i.e., mean total daily "OFF" time at Week 16 – mean total daily "OFF" time at baseline).

The alternative hypothesis will be that μ for Safinamide group is lesser than μ for placebo group. Symbolically, this is expressed as follows:

H_0: μ(safinamide 100 mg/day) = μ(placebo)
H_1: μ(safinamide 100 mg/day) $\neq$ μ(placebo)

An analysis of covariance (ANCOVA) with treatment and centre as independent factor, baseline mean total daily "OFF" time measurement as covariate and the mean change from baseline at Week 16 as dependent variable, with α=0.05 will be used to test this hypothesis.

Results will be reported as Least-Square Means for treatment differences with associated two-tailed 95% confidence intervals and corresponding two-sided p-values. The main time point for comparison between treatment groups is week 16, but other available measurements at other visits will be analyzed with the same model described for the primary analysis.

The following summaries and listings will be provided for diary data:

- The summary of mean change from baseline in mean total daily "OFF" time by scheduled visits (FAS Population)
- The summary of mean change from baseline in mean total daily "OFF" time by scheduled visits (PP Population)
- A by-subject listing of 24-hours diary data
- A by-subject listing of 24-hours total daily "OFF" time

4.9.3 Secondary Efficacy Variables

4.9.3.1 Key Secondary Efficacy Variables

Change from baseline to week 16 in the mean pain severity

The key secondary endpoint is the change from baseline to week 16 in pain severity, as assessed by an 11-point Numerical Rating Scale (NRS) from '0' ("no pain") to '10' ("worst possible pain"). NRS is to be assessed by subjects before going to sleep prior to the baseline visit and the visits at weeks 2, 6, 10 and 16 (visits 3,4,5 and 6), as the worst pain experienced in the past 24 hours (at any time during the day). If the patient suffers for more than one pain, the more severe pain must be recorded.

The null hypothesis for the key secondary endpoint comparison will be that there is no difference between safinamide 100 mg/day and placebo, given as add-on therapy in PD patients with motor fluctuations on stable doses of L-dopa, in change from baseline to week 16 in pain severity, as assessed by an 11-point Numerical Rating Scale (NRS). μ is defined as the mean change from baseline to week 16 in the mean "NRS" by treatment group (i.e., mean "NRS" time at Week 16 – mean "NRS" time at baseline). The alternative hypothesis will be that there is a difference. Symbolically, this is expressed as follows:

H(0): μ(safinamide 100 mg/day) = μ(placebo)
H(1) : μ(safinamide 100 mg/day) $\neq$ μ(placebo)

An ANCOVA model with treatment and centre as independent factor, baseline NRS measurement as covariate and change from baseline as dependent variable, will be used to test this hypothesis (α=0.05).

Results will be reported as Least-Square Means for treatment differences with associated two-tailed 95% confidence intervals and corresponding two-sided p-values. The main time point for comparison between treatment groups is week 16, but other available measurements at other intermediate visits will be analyzed with the same model and only change the definition of dependent variable, which will be considered in an exploratory way.

4.9.3.2 Other Secondary Efficacy Variables

1. Diary Data

In diary, status of "ON" without dyskinesia, "ON" with nontroublesome and "ON" with Troublesome dyskinesia is counted as "ON" status. Following the same analysis for averaged total daily "OFF" time described in primary endpoint, the average total daily "ON" time and "ON" time with no/non-troublesome dyskinesia can be calculated as well.

Change from baseline to week 16 in the mean total daily "ON" time, as assessed by 24-hour patient diary cards

The hypothesis of superiority of safinamide compared to placebo will be assessed using an ANCOVA model parameterized as above (i.e. as for the primary and key secondary endpoint, with the only difference that "baseline mean total daily "ON" time" will be used as covariate), with point estimate of treatment differences reported as Least-Square Mean with associated two-tailed 95% confidence intervals and two-sided p-value. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

The following summaries and listings will be provided for diary data:

- The summary of change from baseline in mean total daily "ON" time by scheduled visits (FAS Population)
- The summary of change from baseline of mean total daily "ON" time by scheduled visits (PP Population)

Change from baseline to week 16 in the mean daily "ON" time with no/non-troublesome dyskinesia, as assessed by 24-hour patient diary cards

The hypothesis of superiority of safinamide compared to placebo will be assessed using an ANCOVA model parameterized as above (using "baseline mean total daily "ON" time with no/non-troublesome dyskinesia" as covariate), with point estimate of treatment differences reported as Least-Square Mean with associated two-tailed 95% confidence intervals and two-sided p-value. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

2. UPDRS Data

The following summaries and listings will be provided for UPDRS:

- The summary of mean change from baseline of UPDRS endpoints by scheduled visits (FAS Population, PP Population and LOCF)
- All the UPDRS data will be presented in Data Listing.

The Parts I-III in UPDRS scale contain 44 items, with each item scored on a 5-point scale. Part IV contains 11 questions with a scale ranging from 0 to 23. Thus, the final total score may range from 0 (no disability) to 199 (total disability). The UPDR score is to be assessed by the Investigator at approximately the same time of day as at the baseline visit, if possible, at least 1 hour after the subject has taken their morning dose of safinamide and is in the optimal "ON" state.

Change from baseline to week 16 in the UPDRS total score during the "ON" phase

The hypothesis of superiority of safinamide compared to placebo will be assessed using an ANCOVA model parameterized as above (with the only difference that "baseline UPDRS total score" will be used as covariate), with point estimate of treatment differences reported as Least-Square Mean with associated two-tailed 95% confidence intervals and two-sided p-value. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

Change from baseline to week 16 in the UPDRS part II (ADL) score during the "ON" phase

The hypothesis of superiority of safinamide compared to placebo will be assessed using an ANCOVA model parameterized as above (with the only difference that "baseline UPDRS part II score" will be used as covariate), with point estimate of treatment differences reported as Least-Square Mean with associated two-tailed 95% confidence intervals and two-sided p-value. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

Change from baseline to week 16 in the UPDRS part III (motor function) score during the "ON" phase

The hypothesis of superiority of safinamide compared to placebo will be assessed using an ANCOVA model parameterized as above (with the only difference that "baseline UPDRS part III total score" will be used as covariate), with point estimate of treatment differences reported as Least-Square Mean with associated two-tailed 95% confidence intervals and two-sided p-value. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

3. CGI-S and CGI-C Data

The CGI-S scale measures global severity of illness at a given point in time. It will be rated on a 7-point Likert-type scale ranging from 1 (normal, not ill at all) to 7 (extremely severe). The CGI-C scale will measure the change in the patient's clinical status from baseline using a 7-point scale, ranging from 1 (very much improved) to 7 (very much worse), with a score of 4 indicating no change. The CGI scores are to be assessed by the Investigator at approximately the same time of

day as at the baseline visit, if possible, at least 1 hour after the subject has taken their morning dose of safinamide and is in the optimal "ON" state.

The following summaries and listings will be provided for CGI-S and CGI-C:

- The summary of CGI-S by scheduled visits (FAS Population, PP Population and LOCF)
- The summary of change from baseline of CGI-C by scheduled visits (FAS Population, PP Population and LOCF)
- All the CGI-S and CGI-C data will be presented in Data Listing.

CGI-S score assessed at week 16

The hypothesis of superiority of safinamide compared to placebo will be assessed using the Wilcoxon-Mann-Whitney test stratified by centre (i.e. Van-Elteren test), whilst point estimate of treatment differences will be reported as Hodges-Lehmann estimators together with associated two-sided nonparametric 95% confidence intervals. The primary conclusion will be based on stratified Wilcoxon-Mann-Whitney test, the Hodges-Lehmann estimators along with 95% confidence interval is only provided as supplemental information.

Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

Change from baseline to week 16 in the CGI-C

CGI-C is collected as change score when compared to Baseline visit. Thus, CGI-C is not collected at Baseline visit. The endpoint of change from baseline to week 16 will be summarized descriptively by using CGI-C data at Week 16 for each treatment groups.

The hypothesis of superiority of safinamide compared to placebo will be assessed using the Wilcoxon-Mann-Whitney test stratified by centre (i.e. Van-Elteren test), whilst point estimate of treatment differences will be reported as Hodges-Lehmann estimators together with associated two-sided nonparametric 95% confidence intervals. The primary conclusion will be based on stratified Wilcoxon-Mann-Whitney test, the Hodges-Lehmann estimators along with 95% confidence interval is only provided as supplemental information.

Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

4. PDQ-39 Data

The PDQ-39 can be summarized into eight dimensions:

1. Mobility: Q1 to Q10.

2. Activities of daily living: Q11 to Q16.

3. Emotional well being: Q17 to Q22.
4. Stigma: Q23 to Q26.
5. Social support: Q27 to Q29.
6. Cognition: Q30 to Q33.
7. Communication: Q34 to Q36.
8. Bodily pain: Q37 to Q39.

Each dimension is coded on a scale from 0 (perfect health, i.e. no problem at all) to 100 (worst health, i.e. maximum level of problem) with each scale being calculated as follows:

Dimension score = the total of the raw scores of each item in the dimension divided by the maximum possible raw score of all the items in the dimension multiplied by 100.

If the response to a question is missing, no scale score is calculated for that individual for that domain. This will preclude calculation of the PDQ-39 single index score from the eight domain scores.

Items are grouped into eight scales that are scored by expressing summed item scores as a percentage score ranging between 0 and 100 (100 = more health problems). A summary index score can subsequently be calculated from the eight domains outlined above. The PDQ-39 summary index is derived by the sum of the eight PDQ-39 scale scores divided by eight (the number of scales), which yields a score between 0 and 100 (100 = more health problems). This is equivalent to expressing the sum of all 39 item responses as a percentage score.

Formula for scoring each domain = $\frac{sum\ of\ scores\ of\ each\ question\ in\ domain}{4\ (max.score\ per\ question) \times nos.questions\ in\ domain} \times 100$

The score under each domain can be calculated as following:

1. Mobility = (scores of questions 1+2+3+4+5+6+7+8+9+10) / ((4 x 10) x 100)
2. Activities of daily living = (scores of questions 11+12+13+14+15+16) / ((4 x 6) x 100)
3. Emotional well-being = (scores of questions 17+18+19+20+21+22) / ((4 x 6) x 100)
4. Stigma = (scores of questions 23+24+25+26) / ((4 x 4) x 100)
5. Social support = (scores of questions 27+28+29) / ((4 x 3) x 100)

   Note: if respondents indicate that they do not have a spouse or partner on question 28 then social support can be calculated as follows:
6. Social support = (scores of questions 27+29) / ((4 x 2) x 100)
7. Cognitions = (scores of questions 30+31+32+33) / ((4 x 4) x 100)
8. Communication = (scores of questions 34+35+36) / ((4 x 3) x 100)

9. Bodily pain= (scores of questions 37+38+39) / ((4 x 3) x 100)
10. Scoring for single index (PDQ-39-SI): single index = Sum of domain scores / 8

The single index represents the overall summary index of Parkinson's Disease.

The PDQ-39 score is to be completed by Investigator at approximately the same time of day as at the baseline visit, if possible, at least 1 hour after the subject has taken their morning dose of safinamide and is in the optimal "ON" state.

Change from baseline to week 16 in the PDQ-39 score

The hypothesis of superiority of safinamide compared to placebo in PDQ-39 dimension scores will be assessed using an ANCOVA model parameterized as the model for primary endpoint (i.e. with treatment and centre as independent factor, baseline mean PDQ-39 score as covariate and change from baseline at week 16 in PDQ-39 scores as dependent variable), with point estimate of treatment differences reported as Least-Square Mean with associated two-tailed 95% confidence intervals and two-sided p-value.

The PDQ-39 summary table will be repeated for summary index, Mobility, Activities of daily living, Emotional well being, Stigma, Social support, Cognitions, Communication, Bodily discomfort. The baseline score under each specific domain is regarded as covariate in ANCOVA model. Analysis results within each treatment group or between treatment groups obtained at intermediate visits (Visit 3- Week 2, Visit 4- Week 6, Visit 5- Week 10) will be considered in an exploratory way.

4.10 Safety Evaluation

All safety summaries and analyses will be based upon the Safety Population as defined in Section 4.5.4.

4.10.1 Adverse Events

AEs will be coded by SOC and PT, using the Medical Dictionary for Regulatory Activities (MedDRA, version 22.0).
AEs will be classified as pre-treatment AEs (PTAEs) and TEAEs, according to the period of occurrence, as follows:

- PTAEs: all AEs occurring before the first dose of IMP and not worsening after the first dose of IMP
- TEAEs: all AEs with an onset date on or after the first dose of IMP or AEs occurred before the first dose of IMP got worsening after the first dose of IMP. All AEs occurring after last dose of IMP and within the safety reporting period are TEAE.

Generally, there will be no imputation of missing values and only observed safety data will be included in the analyses. Where dates are missing or partially missing, adverse events will be assumed to be treatment-emergent, unless there is clear evidence (through comparison of partial dates) to suggest that the adverse event started prior to the first dose of IMP. In the AEs analysis, when relationship to study drug is missing for a TEAE it will be imputed to be drug related.

Individual data listings will include all AEs recorded; individual PTAEs and TEAEs will be listed in subject data listings. No summary table will be provided for PTAEs.

Serious adverse event (SAEs) occurring or worsening after the first dose of IMP are defined as TESAEs.

The number and the percentage of patients reporting TEAEs, treatment emergent SAEs, severe TEAEs, TEAEs leading to discontinuation and TEAEs leading to death will be presented by treatment group, along with the number of events occurring. Only TEAEs will be included in the summary tables. Individual data listings will include all AEs recorded; a separate listing will be provided for treatment-emergent SAEs.

For all TEAE tables, a subject will be counted for most severe event only once for each SOC and each PT, even if the subject reported more than one event under each subcategory. The total number of events documented per SOC and PT will also be displayed. All TEAEs will be summarized by treatment group and overall. Unless specified otherwise, TEAE tables will be ordered in terms of decreasing number of subjects for SOC and then PT within the SOC in the Safinamide group then by Placebo group, and then alphabetically for SOC and PT within the SOC if the frequency is tied.

For each subject and each adverse event, the severity recorded will be attributed and used in the by-severity summaries. Similarly, the relationship (related to treatment) will be attributed and used in the by- relationship summaries.

An overview table will summarize the number and percentage on subject level of the following categories:

- Any TEAE
- Any TEAE related to IMP
- Any severe TEAE
- Any severe TEAE, related to IMP
- Any TEAE with outcome of death
- Any TEAE with outcome of death, related to IMP
- Any TESAE
- Any TESAE, related to IMP
- Any TEAE leading to study discontinuation
- Any TEAE leading to discontinuation of IMP
- Any TESAE leading to study discontinuation
- Any TESAE leading to discontinuation of IMP
- Any TEAE leading to discontinuation of IMP, related to IMP
- Any TEAE leading to IMP reduced

TEAE summaries will present the number and percentage of subjects reporting at least one TEAE. The following summaries will be provided:

- TEAEs by SOC and PT
- TEAEs by SOC and PT, by severity
- TEAEs by SOC and PT, by relationship
- TEAEs related to IMP, by SOC and PT
- TESAEs by SOC and PT
- TESAEs by SOC and PT, by severity
- TESAEs by SOC and PT, by relationship
- TESAEs related to IMP, by SOC, PT
- TEAE leading to study discontinuation, by SOC and PT
- TEAE leading to discontinuation of IMP, by SOC and PT
- Any TESAE leading to study discontinuation
- Any TESAE leading to discontinuation of IMP
- TEAE leading to death, by SOC and PT

The number and percentage of deaths will also be summarized.

The following by-subject AE listings will be provided:

- AE
- PTAEs
- TEAEs
- TESAEs
- TEAEs related to IMP
- TEAEs leading to Discontinuation of IMP
- TEAEs leading to drug permanently stopped
- TEAEs leading to death

The listings will include: subject identifier, age, sex, adverse event (SOC, PT, and verbatim term), date of onset, date of resolution, severity, seriousness, action taken, outcome and relationship.

4.10.2 Clinical Laboratory Evaluation

Laboratory values (hematology, chemistry and pregnancy test) will be listed by subject and study time point for continuous variables. The baseline for the laboratory values will be the results obtained on the last measurement before the first dose.

All values outside the clinical reference ranges will be flagged in the data listings. The abnormal values will be flagged with 'L' for values below the lower limit of the clinical reference range and 'H' for values above the upper limit of the clinical reference range and included in the listings. The Investigator will assess whether the values outside the clinical reference range are clinically

significant and these will be reported as "abnormal, NCS" or "abnormal, CS". Clinically significant laboratory values will be recorded by the Investigator as AEs.

Quantitative laboratory assessment reported as "<X", i.e. BLQ, or ">X", i.e. above the upper limit of quantification, will be converted to X for the purpose of quantitative summaries, but will be presented as recorded, i.e. as "<X" or ">X" in the listings.

Haematology and clinical chemistry results at Visit 1 and Visit 6 will be converted to standard international units then summarized by treatment group descriptively for continuous variables. The following summaries will be provided for laboratory data:

- Mean laboratory value by visits and change from baseline for continuous data at Visit 6 are presented by treatment groups and overall.
- The number and percentage of subjects with shift lab results from baseline to Visit 6 is presented by category of "Normal", "abnormal NCS", and "abnormal CS", and by treatment groups.

The following laboratory data listings will be provided:

- By-subject listings of laboratory data will be provided with investigator's evaluations
- By-subject listings of Pregnancy Test

4.10.3 Vital Signs, Body Weight, 12-Lead Electrocardiogram (ECG), and Physical Examinations (PE)

Vital sign values by visits and change from baseline of vital signs parameters and body weight will be presented with descriptive statistics by scheduled time point, treatment group and overall.

The following summaries and listings will be provided for vital signs:

- A summary of vital signs (weight, heart rate and systolic and diastolic blood pleasure) by treatment group and overall.
- By-subject listings of vital signs and body weight data will be provided with investigator's evaluations.

The following summaries will be provided for ECG parameters and PE:

- The number and percentage of subjects with shift ECG evaluations from baseline to Visit 6 is presented by category of "Normal", "abnormal NCS", and "abnormal CS", and by treatment groups.
- The number and percentage of subjects with PE as "Normal", "abnormal NCS", and "abnormal CS", by scheduled time point, body system treatment group and overall.
- By-subject listings of ECG data will be provided with investigator's evaluations.
- By-subject listings of PE data will be provided with investigator's interpretation.

4.11 Determination of Sample Size

Sample size was computed through a Monte Carlo study with 1000 runs and using a Fixed Sequence Procedure (1) to account for multiplicity over the study primary endpoint (change from baseline to week 16 in the mean total daily "OFF" time) and the key secondary endpoint (change from baseline to week 16 in pain severity, as assessed by an 11-point Numerical Rating Scale – NRS). The Fixed Sequence Procedure implies that the key secondary endpoint will be testable provided that the primary endpoint has achieved statistical significance with a two-tailed p-value ≤ 0.05.

Based on the Monte Carlo simulation it was estimated that a total sample size of 260 patients (130 in the safinamide and 130 in the placebo groups) ensures 90% power to detect a mean difference in the "OFF" time at least 0.9 h between the safinamide and placebo groups with a two-sided significance level (alpha) of 0.05 using a two-sample t-test and assuming standard deviations of 2.35 for safinamide and 2.06 for placebo. Effect size and standard deviation estimates used in sample size computations are gathered from the Settle Statistical Report (Table 15.2.13).

Moreover, the same Monte Carlo study showed that the total sample size of 260 patients would also permit a marginal power equal to 88% to detect 1 point treatment difference in the NRS between the safinamide and placebo groups with a two-sided significant level (alpha) of 0.05 using a two-sample t-test and assuming a pooled standard deviations of 2.0. Standard deviation was estimated on the basis of previous post-hoc analyses of pain (1) whilst a 1-point treatment difference is considered to be a clinically meaningful treatment effect (2).

Assuming an attrition rate equal to 15% a total of approximately 306 patients will be randomized (153 in the safinamide and 153 in the placebo groups). The sample size calculations were performed using the Mediana package (3).

4.12 Changes in the Conduct of the Study or Planned Analysis

According to protocol 20.4.8, descriptive statistics for PE data at each visit will be presented overall and by treatment. Since only categorical data are collected for PE, relevant PE analysis as categorical data is updated in SAP section 4.10.3.

According to protocol 10.1, the duration of Parkinson's Disease should be reported as part of demographic information. The diagnosis date of study indication was not collected in CRF, therefore, this variable is not able to be summarized for this study.

5 REFERENCES

1. Maurer, Hothorn, Lehmacher. Multiple comparisons in drug clinical trials and preclinical assays: a-priori ordered hypotheses. In Biometrie in der chemisch-pharmazeutischen Industrie. Vollmar J (Editor). Stuttgart: Fischer Verlag. 1995; 6: 3-18.
2. Rascol O, Zesiewicz T, Chaudhuri KR, Asgharnejad M, Surmann E, Dohin E et al. A randomized controlled exploratory pilot study to evaluate the effect of rotigotine

transdermal patch on Parkinson's Disease-associated chronic pain. J Clin Pharmacol 2016; 56(7): 852-861.
3. Dmitrienko, A., Pulkstenis, E. Clinical Trial Optimization Using R. Chapman & Hall/CRC Biostatistics Series, 2017; 4-28.

5.1 APPENDIX 1: STUDY FLOW CHART

Study Period	Screening period	Baseline	Treatment period			End of Treatment (EOT) / Early Termination (ET)[1]	Telephone follow-up	Unscheduled visit
Week	**-2 to 0**	**0**	**2**	**6**	**10**	**16**	**17**	
Day	**-14 to -2**	**1**	**14 ±3**	**42 ±3**	**70 ±3**	**112 ±3**	**119 ±3**	
Visit	**1**	**2**	**3**	**4**	**5**	**6**	**7**	
Informed consent	X							
Eligibility criteria	X	X						
Randomization		X						
Demographics	X							
Medical history/diagnosis	X							
Hoehn & Yahr staging	X							
Neurological examination[2]	X							
Prior and concomitant medications[3]	X	X	X	X	X	X	X	X
Vital signs	X	X	X	X	X	X		X
12-lead ECG	X					X		X
Physical examination	X					X		X
Diary Issue and Training[4]	X	X	X	X	X			
Diary Review		X	X	X	X	X		
UPDRS [5]		X	X	X	X	X		
CGI-S[5]		X	X	X	X	X		
CGI-C[5]			X	X	X	X		
PDQ-39[5]		X	X	X	X	X		
NRS[6]		X	X	X	X	X		
Laboratory exams	X					X		
Pregnancy test (urine)[7]	X					X		
Adverse events[8]	X	X	X	X	X	X	X	X
Dispense randomised medication		X[9]	X	X	X			X
Drug accountability			X	X	X	X		X

ECG: electrocardiogram; UPDRS: Unified Parkinson's Disease Rating Scale; CGI-S: Clinical Global Impression-Severity; CGI-C: Clinical Global Impression-Change; PDQ-39: Parkinson's Disease Questionnaire-39 items; NRS: Numerical Rating Scale.

1. Subjects who prematurely withdraw from the study while receiving study medication should complete the EOT Visit assessments, when possible.
2. Neurological examination at screening visit only.
3. Details of excluded and permitted concomitant treatments are presented in Section 11.0 of the protocol.
4. The diary should be completed by the subject two days before the baseline visit and the visits at weeks 2, 6, 10 and 16 (visits 3,4,5 and 6) for recording of "OFF" and "ON" time
5. To be evaluated at approximately the same time of day as at the baseline visit, if possible at least 1 hour after the subject has taken their morning dose of safinamide and is in the optimal "ON" state.
6. NRS will record PD pain intensity in the evening before going to sleep prior to the baseline visit and the visits at weeks 2, 6, 10 and 16 (visits 3,4,5 and 6), as the worst pain experienced in the past 24 hours (at any time during the day). If the patient suffers for more than one pain, the more severe pain must be recorded.
7. All women of child-bearing potential.
8. The Investigator must report all AEs which occur during the study following written informed consent.
9. The first dose of safinamide or placebo (50 mg once daily) will be administered at the study centre. The dose of safinamide or placebo will be titrated the day after the Visit 3/week 2, ideally at day 15 at home to 100 mg once daily. See Section 7.0 of the protocol.

5.2 APPENDIX 2: Partial Date Conventions

Imputed dates will NOT be presented in the listings. The following rules in the given table will be applied for each case. For partially missing stop date of medication, please include last known assessment date (LKAD) as a reference timepoint to impute missing date. This date is patient's last date in study, ideally is equal to end of study date, or the maximum AE end date. However, LKAD should be determined on the basis of the real data.

5.2.1 Algorithm for Prior/Concomitant Medications, Therapies or Procedures:

Start Date	Stop Date	Action
Known	Known	If stop date is prior to the date of first dose of study drug, considered as prior; if start date is on or after the date of first dose of study drug, considered as concomitant. If start date is prior to the date of first dose of study drug and stop after first dosing date, considered as concomitant.
	Partially Missing	**If the day and month of stop date is missing:** • If year of partial date < year of the LKAD, assigned the last day of the month and the last month (i.e. December 31 st). • If year of partial date = year of the LKAD, assign the LKAD date to the missing fields. **If only the day of stop date is missing:** • If month of partial date < month of the LKAD, assign the last day of the month. • If month of partial date = month of the LKAD, assign the LKAD date to the missing fields. If the imputed stop date is prior to the date of first dose of study drug, considered as prior; else considered as concomitant. **Noted: If year of partial date is missing, please considered this as "Concomitant" Medications, Therapies or Procedures.**
	Completely Missing	Considered as concomitant.
Partially Missing	Known	The first day of the month and January will be used if the start day/month is missing. If stop date is prior to the date of first dose of study drug, considered as prior; else considered as concomitant.

Start Date	Stop Date	Action
	Partially Missing	The first day of the month and January will be used if the start day and month is missing. **If the day and month of stop date is missing:** • If year of partial date < year of LKAD, assigned the last day of the month and the last month (i.e. December 31 st). • If year of partial date = year of the LKAD, assign the LKAD date to the missing fields. **If only the day of stop date is missing:** • If month of partial date < month of the LKAD, assign the last day of the month. • If month of partial date = month of the LKAD, assign the LKAD date to the missing fields. If the imputed stop date is prior to the date of first dose of study drug, considered as prior; else considered as concomitant. **Noted: If year of partial date is missing, please considered this as "Concomitant" Medications, Therapies or Procedures.**
	Completely Missing	The first day of the month and January will be used if the start day/month is missing. Considered as concomitant.
Completely Missing	Known	If stop date is prior to the date of first dose of study drug, considered as prior; else considered as concomitant.
	Partially Missing	If the day and month of stop date is missing: • If year of partial date < year of LKAD, assigned the last day of the month and the last month (i.e. December 31 st). • If year of partial date = year of the LKAD, assign the LKAD date to the missing fields. **If only the day of stop date is missing:** • If month of partial date < month of the LKAD, assign the last day of the month. • If month of partial date = month of the LKAD, assign the LKAD date to the missing fields. If the imputed stop date is prior to the date of first dose of study drug, considered as prior; else considered as concomitant. **Noted: If year of partial date is missing, please considered this as "Concomitant" Medications, Therapies or Procedures.**
	Completely Missing	Considered as concomitant.

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 (liuy25)
Title: PPD
Date: Tuesday, 09 November 2021, 05:08 AM GMT Standard Time
Meaning: Document contents approved.
==========