A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Examine the Efficacy, Safety and Tolerability of APL-130277 in Levodopa Responsive Patients with Parkinson's Disease Complicated by Motor Fluctuations ("OFF" Episodes).

Study code: CTH-300

Phase III study

STATISTICAL ANALYSIS PLAN

Signatures:

Statistical Analysis Plan was prepared by:

[redacted]
Consulting Statistician, Clintrex

7 NOV 2017
Date

[redacted]
Associate Director, Biostatistics, Sunovion

2 NOV 2017
Date

Statistical Analysis Plan was reviewed/approved by:

[redacted]
Date: 2017.11.02 23:03:58 -04'00'
Sr. Director, Clinical Development, Sunovion

Date

[redacted]
Sr. Director, Biostatistics, Sunovion

3 NOV 2017
Date

Table of Contents

1 Abbreviations 5
2 Introduction 7
3 Study objectives 7
4 Design and type of the study 7
5 Endpoints 8
5.1 Primary endpoint 8
5.2 Key secondary endpoint 8
5.3 Other secondary endpoints 9
5.4 Other subject-reported secondary endpoints 11
6 Sample size considerations 12
7 Statistical hypotheses 12
8 Analysis Populations 13
8.1 All Available population 13
8.2 Modified Intention-To-Treat population 13
8.3 Intention-To-Treat population 13
8.4 Safety population 13
8.5 Maintenance Phase Safety population 13
8.6 Completer population 14
8.7 Per Protocol population 14
9 General statistical considerations 14
9.1 Adjustment for covariates 14
9.2 Handling of drop-outs or missing data 15
9.3 Interim analyses and data monitoring 15
9.4 Multiple comparison/multiplicity 15
9.5 Examination of subgroups 16
10 Disposition 17
11 Demographic and other baseline characteristics 18
12 Concomitant medication/treatment 19
13 Medical history 20
14 Analysis of efficacy 21
14.1 Primary endpoint 21
14.1.1 Patient Global Sensitivity analyses of the primary endpoint 22
14.1.2 Graphical summaries of the primary endpoint 23
14.2 Key Secondary endpoint 24
14.2.1 Graphical summaries of percent of patients with patient-rated full “ON” 25
14.3 Other secondary endpoints 25
14.4 Other patient-reported endpoints 27
15 Analysis of safety and tolerability 27

15.1 Extent of exposure 27
15.2 Compliance 28
15.3 Adverse events 28
15.4 Laboratory safety variables 31
15.5 Vital signs 31
15.6 ECG 32
15.7 Other safety variables 33
16 Other variables 33
17 Deviations from the analyses planned in the study protocol 33
18 Execution of statistical analyses 34
19 Hardware and software 34
20 References 34

1 Abbreviations

AAP	All Available Population
AE	Adverse Event
ANCOVA	Analysis of Covariance
ATC	Anatomical Therapeutic Chemical
BP	Blood Pressure
CGI-S	Clinical Global Impression of Severity
CGI-I	Clinical Global Impression of Improvement
CI	Confidence Interval
C-SSRS	Columbia Suicide Severity Rating Scale
DSMB	Data and Safety Monitoring Board
EOS	End Of Study
EQ-5D	European Quality of Life – 5 Dimensions
ESS	Epworth Sleepiness Scale
HR	Heart Rate
ITT	Intention-To-Treat
L-Dopa	Levodopa
LOCF	Last Observation Carried Forward
LS	Least Square
MAR	Missing At Random
MCMC	Markov Chain Monte Carlo
MDS	Movement Disorders Society
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
mITT	Modified Intention-To-Treat
MMRM	Mixed Model for Repeated Measures
MMSE	Mini–Mental State Examination
MV	Maintenance Visit
PD	Parkinson's Disease

PDQ-39	Parkinson's Disease Questionnaire-39
PGI-I	Patient Global Impression of Improvement
PGI-S	Patient Global Impression of Severity
PMM	Pattern Mixture Model
PT	Preferred Term
QUIP-RS	Questionnaire for Impulsive-Compulsive Disorders in Parkinson's disease – Rating Scale
RR	Respiratory Rate
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation
SEM	Standard Error of the Mean
SI	International System of Units
SOC	System Organ Class
SV	Screening Visit
TEAE	Treatment-Emergent Adverse Event
TV	Titration Visit
UPDRS	Unified Parkinson's disease Rating Scale
VAS	Visual Analog Scale
WHO-DD	World Health Organization Drug Dictionary

2 Introduction

This Statistical Analysis Plan (SAP) describes the statistical analysis for study CTH-300 and was finalized prior to unblinding of the study. The SAP will serve as a compliment to the study protocol and supersedes it in case of differences. In case of major differences between the study protocol and SAP (e.g. changes in the analysis related to the primary endpoint), a protocol amendment will be considered.

3 Study objectives

The primary objective of this study is to evaluate the efficacy and safety of APL-130277 versus placebo in patients with Parkinson's disease (PD) over a 12 week period.

4 Design and type of the study

This is a 12-week, prospective, multi-center, randomized, double-blind, placebo-controlled, Phase 3 study in levodopa (L-Dopa) responsive PD patients with motor fluctuations, designed to determine the efficacy, safety and tolerability of APL-130277.

The study includes initial Screening Visits, followed by an initial Dose Titration Phase in which individual responses to single doses of APL-130277 are evaluated in order to determine the optimal dose of APL-130277 for treating "OFF" episodes to be used in the maintenance phase of the study. Once complete, patients will be randomized to either the APL-130277 or placebo (ratio 1:1) and begin the Maintenance Treatment Phase of the study, where they will self-administer study medication in up to 5 doses per day to treat "OFF" episodes over the course of 12 week study in the at-home portion of the study. Patients will return to the clinic at regular intervals for safety and efficacy assessments (including in-office assessments of the primary endpoint). Overall duration of participation will be approximately 135 days.

The following visits will be performed:

- Screening Visits (SV1 and SV2)
- Telephone Call (T1)
- Dose Titration Phase
    - Titration Visit 1 (TV1)
    - Titration Visit 2 (TV2)
    - Titration Visit 3 (TV3)
    - Titration Visit 4 (TV4)
    - Titration Visit 5 (TV5)

- Titration Visit 6 (TV6)

- Maintenance Treatment Phase
  - Maintenance Visit 1 (MV1)
  - Maintenance Visit 2 (MV2)
  - Maintenance Visit 3 (MV3)
  - Maintenance Visit 4 (MV4)
  - Telephone Call (T2, T3, T4)
  - End of Study Visit (EOS)
  - Unscheduled Dose Adjustment Visits.

The patients will be randomized centrally at a study level using a computer-generated randomization code. No stratification factors will be used.

5 Endpoints

5.1 Primary endpoint

The primary endpoint of the study is the mean change from pre-dose in MDS-UPDRS MOTOR score at 30 minutes after dosing at the 12 week visit (MV4) of the Maintenance Treatment Phase.

The MDS-UPDRS MOTOR score will be calculated as the sum of the individual items of the MDS-UPDRS Part III questionnaire (items 3.1 – 3.18) and will be obtained separately at each assessment time point. Missing individual items will be imputed using the 2 non-missing values at time points adjacent to the missing item on the same date. The maximum of the 2 adjacent values will be assigned as the score for the missing individual item. However, pre-dose values will not be assigned as post-dose values and if one of the adjacent values for a post-dose value is a pre-dose value, only 1 adjacent value will be used. If a pre-dose value is missing, the pre-dose value at the prior visit will be used. The maximum of the 2 adjacent values represents the worse outcome out of these potential values and may be seen as a conservative imputation. If there are more than three individual missing items at a given time point, no imputation will be performed and MDS-UPDRS MOTOR score will be assigned as missing. The MDS-UPDRS MOTOR score will be calculated after imputation of the missing item(s) as described above.

5.2 Key secondary endpoint

The key secondary endpoint of the study is the percentage of patients with a patient-rated full “ON” response within 30 minutes after dosing at the 12 week visit (MV4) of the Maintenance Treatment Phase.

A full "ON" response is defined as a period of time where in the judgment of the patient the medication is providing full benefit with regard to mobility, stiffness, slowness and other PD features comparable to or better than that obtained with their standard dose of oral levodopa and other anti-parkinsonian medications prior to beginning the study.

An assessment of the "OFF"/"ON" state will be performed at 0, 15, 30, 45, 60 and 90 minutes after dosing at each visit. Patients will also be asked if they attained a full "ON" state anytime within 30 minutes of dosing. This endpoint will be based on the latter question (did the patient attain a full "ON" state anytime within 30 minutes of dosing). However, if this question has not been completed but the patient reported the "OFF"/"ON" state at 15 and/or 30 minutes, the 15 and 30 minute assessments will be used to derive the endpoint. In this case, if the patient has reported an "ON" state at either of the two time points the patient will be considered as having the full "ON" response within 30 minutes after dosing. The visits at which there is no 15 or 30 minute assessment of the "OFF"/"ON" state and the question about attaining a full "ON" state anytime within 30 minutes of dosing is missing will be set as having a missing value.

5.3 Other secondary endpoints

1. The percentage of instances where a full "ON" response was achieved at 30 minutes after self-administration of study medication based on the home dosing diary entries during the 2 days prior to the 12 week visit (MV4).

   During the Maintenance Treatment Phase of the study, patients will complete the home dosing diary on the 2 days prior to their next scheduled in-clinic visit. The patients will self-administer their doses of randomized treatment (APL-130277 or placebo; 1:1 ratio) in order to treat up to 5 "OFF" episodes per day. The patients will fill in the time when study treatment is self-administered and the "ON"/"OFF" status at 30 minutes following dosing. In total, each patient can record up to 10 episodes at 3 visits after the randomization (MV2, MV3 and MV4), i.e. a total of up to 30 episodes. For each patient, the percentage of episodes in which the full "ON" response was achieved at 30 minutes out of all recorded episodes will be calculated. In case the time of self-administration has been recorded but the corresponding "ON"/"OFF" status is missing, the status will be classified as "OFF". The percentages calculated separately for each patient will be used as response variables in the statistical analysis. For patients who did not record any episodes during the Maintenance Treatment Phase, the endpoint will be set as missing.

2. Mean change from pre-dose in MDS-UPDRS MOTOR score at 15, 45, 60 and 90 minutes at the 12 week visit (MV4) of the Maintenance Treatment Phase.

   These variable will be defined using rules similar to the primary endpoint.

3. Time (in minutes) from dosing to when study medication is starting to provide an "ON" effect at the 12 week visit (MV4) of the Maintenance Treatment Phase.

The time from dosing until the patient reports to have an effect will be used as the endpoint. For patients who do not report an effect during the follow-up period, the data will be censored at 90 minutes.

4. Percent of patients with a patient-rated full "ON" response within 30 minutes, whose duration from time when study medication begins to have an effect until their "OFF" (if applicable) lasts for at least 30 minutes at the 12 week visit (MV4) of the Maintenance Treatment Phase.

   The identification of patients who experienced a full "ON" response within 30 minutes will be defined as identical to that for the key secondary endpoint. Duration of the response will be calculated by evaluating the difference in time from the "Time to when study treatment started to have an effect as per subject assessments" and the "Time to OFF" per patient assessment. If the latter question is left blank or marked as "N/A", time "OFF" will be identified as per the Patient confirmation of "OFF" and "ON" following dosing at 15, 30, 45, 60 and 90 minutes after dosing. If the patient did not turn "OFF" after dosing, duration will be assumed to have ended at 90 minutes after dosing. For a blank value or "N/A" ticked for the question "Time to when study medication started to have an effect as per subject assessment", the data will be treated as having a missing value. Patients who do not turn "ON" or who turn "ON", but with a duration less than 30 minutes will be considered non-responders in this analysis.

5. Clinical Global Impression of Improvement (CGI-I) post dosing.

   The non-missing values will be categorized as improvements (very much improved, much improved, minimally improved) or non-improvements (no change, minimally worse, much worse, very much worse). The missing values will be considered as non-improvements. CGI severity (CGI-S) will also be assessed.

6. Patient Global Impression of Improvement (PGI-I) post dosing.

   The non-missing values will be categorized as improvements (very much improved, much improved, minimally improved) or non-improvements (no change, minimally worse, much worse, very much worse). The missing values will be considered as non-improvements. PGI severity (PGI-S) will also be assessed.

7. Mean change from baseline in Parkinson's Disease Questionnaire-39 (PDQ-39) summary index score.

   Change from SV to MV4 in PDQ-39 sub-scores (mobility score, activities of daily living, bodily discomfort score, emotional wellbeing score, social support score, communication score, cognitive impairment score, and stigma score) and summary index score will be calculated. The questionnaire provides scores on eight dimensions as outlined below:

   - mobility (10 items, #1 to 10)
   - activities of daily living (6 items, #11 to 16)
   - emotional well-being (6 items, #17 to 22)

- stigma (4 items, #23 to 26)
- social support (3 items, #27 to 29)
- cognitions (4 items, #30 to 33)
- communication (3 items, #34 to 36)
- bodily discomfort (3 items, #37 to 39).

Items are scored from 0 (never) to 4 (always). Dimension scores are obtained by dividing the sum of the item scores by the maximum possible score for any given dimension and expressing this as a percentage. For example:

- mobility = (sum of scores of #1 to 10)/(4 x 10) x 100
- activities of daily living = (sum of scores of #11 to 16)/(4 x 6) x 100.

For social support, if the response indicates that a patient does not have a spouse or partner for #28, social support can be calculated as [(sum of scores of #27 and 29)/(4 x 2) x 100].

A summary index is then calculated as the sum of the total score of the dimensions divided by the number of dimensions, i.e. (sum of dimension scores / 8). If any item score is missing, the relevant dimension score and the summary index will be missing.

8. Change from baseline in MDS-UPDRS – Part II: Motor Aspects of Experiences of Daily Living.

   The change from SV to MV4 in MDS-UPDRS Part II score will be evaluated. The MDS-UPDRS Part II score will be calculated as the sum of the individual items of the MDS-UPDRS Part II questionnaire (items 2.1 – 2.13). Missing individual items will not be imputed. If there is at least 1 missing item, the corresponding MDS-UPDRS Part II score will be set as missing.

9. Evaluation of safety and tolerability data collected, including 12-lead ECGs, orthostatic hypotension, oropharyngeal and dopaminergic AEs.

   The evaluation of the safety and tolerability data is described in section 15 of this document.

5.4 Other subject-reported secondary endpoints

1. Change in sleep measures on the Epworth Sleepiness Scale (ESS).

   The ESS is used to determine the level of daytime sleepiness and evaluated as change from SV to MV4. There are 8 situations listed for which patients rate their likelihood of dozing or sleeping (0=no chance of dozing, 1=slight chance of dozing, 2=moderate chance of dozing, and 3=high chance of dozing). The total score is the sum of 8 item scores and can range between 0 and 24. In case of missing item scores, the missing value will be replaced by the average of non-missing scores at the same visit from the same patient. In case all item scores are missing, the total score will be set as missing. The higher total score indicates the higher level of daytime sleepiness. A score of 10 or more is considered sleepy, and a score of 18 or more is very sleepy.

2. Change in European Quality of Life – 5 Dimensions (EQ-5D).

   The EQ-5D is a utility scale consisting of three components: health state dimensions, health state thermometer scale and health state index.

   The health state dimensions will be described by the 5 dimensions of the EQ-5D (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension has 5 response choices, listed in order of increasing severity. The health state dimensions will be evaluated by presenting the distribution of responses separately for each of the 5 dimensions.

   The health state thermometer scale asks respondents to rate their present health status on a 0 to 100 visual analog scale (VAS). The change from SV to MV4 on the VAS scores will be evaluated.

The health state index score will be calculated based on EQ-5D-5L Crosswalk Index Value Calculator - EuroQol, using the dimension scores from the 5 dimensions, ranging between 1.0 (best imaginable health) and -0.594 (worst imaginable health).

6 Sample size considerations

The study intends to show the superiority of APL-130277 compared with placebo. Assuming a treatment difference of 7 points and standard deviation of 10 points, a sample size of 44 subjects per group or 88 patients for two groups will provide ≥ 90% power to detect a statistically significant difference at the 0.05 level (2-sided), using a two-sample t-test. Taking into consideration a 10% dropout rate during the titration phase and a 15% dropout rate during the maintenance phase, the study plans to enroll approximately 126 patients into the Dose Titration Phase and to randomize approximately 114 patients into the Maintenance Treatment Phase.

7 Statistical hypotheses

The primary objective of this study is to show that APL-130277 is superior to placebo in improving the motor function, assessed as the mean change from pre-dose in Movement Disorders Society (MDS) Unified Parkinson's Disease Rating Scale (UPDRS) Part III Motor Examination (MDS-UPDRS MOTOR) score at 30 minutes after dosing at the 12 week visit (MV4) of the Maintenance Treatment Phase. That is, the null-hypothesis to be tested is

H_0: APL-130277 is the same as placebo in its effect on the motor function

against the two-sided alternative

H_1: Either of the treatment groups is superior to the other in its effect on the motor function.

8 Analysis Populations

The following analysis populations will be used for the analysis of the disposition and efficacy and safety data.

8.1 All Available population

The All Available Population (AAP) will include all patients who have consented for the study, including screening failures. Unless specified otherwise, the AAP will be used for patient listings and for the summary of patient disposition.

8.2 Modified Intention-To-Treat population

All patients who are randomized and receive at least one post-randomization dose of study medication (APL-130277 or placebo) will comprise the modified Intention-To-Treat (mITT) population. The mITT population will be used for the efficacy analysis, and patients will be grouped according to the randomized treatment group.

8.3 Intention-To-Treat population

All patients who are randomized will comprise the Intention-To-Treat (ITT) population and patients will be grouped according to the randomized treatment group.

8.4 Safety population

All patients who are enrolled and receive at least one dose of APL-130277 during the Dose Titration Phase will be included in the safety population. The Safety population will be used for the analysis of the safety endpoints from the Dose Titration Phase and for the pooled data from Dose Titration Phase and Maintenance Treatment Phase.

8.5 Maintenance Phase Safety population

All patients who receive at least one dose of study medication (APL-130277 or placebo) during the Maintenance Treatment Phase will comprise the Maintenance Phase Safety population. The Maintenance Phase Safety population will be used for the analysis of the safety endpoints from the Maintenance Treatment Phase. The patients will be grouped according to the medication actually received (APL-130277 or placebo) during the Maintenance Phase. In cases where the patient received at least one dose of both APL-130277 and placebo, they will be grouped into the APL-130277 arm for safety analysis.

8.6 Completer population

All mITT patients who have a valid MDS-UPDRS MOTOR score at pre-dose and after 30 minutes post-dose at baseline and MV4.

8.7 Per Protocol population

All mITT patients who complete the study with no major protocol deviation.

9 General statistical considerations

All data from all patients entered into the database will be included in patient data listings. The listings will be generally sorted by center and patient number (and by visit and by time point, if applicable), unless specified otherwise.

All applicable data will be summarized separately for the Dose Titration Phase, Maintenance Treatment Phase and for the pooled data from the Dose Titration Phase and Maintenance Treatment Phase. The Maintenance Treatment Phase data will be summarized by treatment group (APL-130277 or placebo). Selected analyses will be performed based on the dose required to achieve a full "ON" response (see Section 9.5 for more details). Where appropriate, data will be summarized by visit and/or time point. Screening visit (SV) assessments may be performed at either SV1 or SV2. If performed at both visits, the latter assessment will be used as the SV assessment for the purposes of analysis. Unscheduled or repeat assessments will not be included in summary tables, but will be included in listings.

Continuous variables will be summarized using the number of observations (n), mean, standard deviation (SD), median, minimum, and maximum. Standard error of the mean (SEM) will also be provided for summaries of efficacy data, if relevant.

Descriptive statistics for categorical data will include frequency counts and percents. The total number of patients in the treatment group (N) will be used as the denominator for percent calculations, unless stated otherwise.

Significance testing will be 2-sided using $\alpha = 0.05$, unless otherwise specified.

All statistical analyses and summaries will be produced using SAS version 9.3 or higher. Deviations from the statistical plan will be reported in the clinical study report, including the rationale for use.

9.1 Adjustment for covariates

For the analysis of efficacy endpoints, the baseline value of the endpoint in question will generally be included as a baseline covariate (for continuous endpoints) or as a stratification factor (for categorical endpoints), if available. Use of covariates for statistical adjustment is discussed below where relevant.

9.2 Handling of drop-outs or missing data

Several different methods to handle the missing data in efficacy assessments will be used.

- For the primary efficacy analysis, likelihood-based modeling approach will be used to handle incomplete data. For this purpose, Mixed Model for Repeated Measures (MMRM) will be applied, see Section 14.1.
- Sensitivity analysis for primary efficacy data will be conducted using the Multiple Imputation (MI) approach, i.e. by replacing each missing value with a set of plausible values that represent the uncertainty about the right value to impute. LOCF will also be used for sensitivity analysis (see Section 14.1).

Safety data will not be subject to any imputation and will be summarized on an observed case basis.

9.3 Interim analyses and data monitoring

No interim analyses of the efficacy data are foreseen. The safety data will be monitored during the study conduct by an independent Data and Safety Monitoring Board (DSMB) to determine if safety monitoring can be modified. The responsibilities of the DSMB include:

- Monitor safety and study conduct;
- On a regular basis, review unblinded summaries of demographic, patient disposition, study conduct and safety data (provided by an independent statistician);
- Evaluate results of an analysis performed after 50% patients have completed the Dose Titration Phase to determine if it is safe and appropriate to reduce the extent of safety evaluations during titration, and to make recommendations for amending the protocol.

The full responsibilities and purview of the DSMB are outlined in the DSMB Charter , which was approved by the Sponsor prior to the implementation of any DSMB review (see Section 17.4.7 of the Protocol).

9.4 Multiple comparison/multiplicity

The multiplicity due to multiple efficacy endpoints will be addressed by using a hierarchical testing approach.

The primary and secondary endpoints will be tested in a fixed sequential manner. The order of the testing is outlined below. See Section 5 for the definition of the primary and secondary endpoints.

1. Primary endpoint: Mean change from pre-dose in MDS-UPDRS MOTOR score after 30 minutes at MV4
2. Key secondary endpoint: percentage of patients with a patient-rated full “ON” response within 30 minutes at MV4

3. Percent of patients with a patient-rated full "ON" response within 30 minutes, whose duration from time when study medication begins to have an effect until their "OFF" (if applicable) lasts for at least 30 minutes at MV4
4. Patient Global Impression of Improvement (PGI): The percentage of patients improved (ie, very much improved, much improved or minimally improved) at MV4
5. Clinician Global Impression of Improvement (CGI): The percentage of patients improved (i.e., very much improved, much improved or minimally improved) at MV4
6. Mean change from SV to MV4 in MDS-UPDRS – Part II: Motor Aspects of Experiences of Daily Living
7. The percentage of instances where a full "ON" response was achieved at 30 minutes after self-administration of study treatment in the outpatient setting based on the home dosing diary entries during the 2 days prior to MV4
8. Mean change from SV to MV4 in PDQ-39 summary index score
9. Mean change from pre-dose in MDS-UPDRS MOTOR score at 15 minutes at MV4
10. Time (in minutes) to when study medication is starting to have an effect at MV4

First, the endpoint ranked as first will be tested and the difference will be declared statistically significant if the nominal p-value is less than 0.05. Second, in case that the difference for the first endpoint is statistically significant, the endpoint ranked as second will be tested, and the difference will be declared statistically significant if the nominal p-value is less than 0.05. The testing will continue as long as the previously ranked endpoint was statistically significant.

In case all endpoints are not evaluated as a part of the hierarchical testing procedure, the statistical tests will be reported in any case and interpreted in a descriptive manner

9.5 Examination of subgroups

At least the following subgroup analyses have been pre-planned. The subgroup analyses are considered exploratory and will be performed for selected efficacy endpoints (at least the primary endpoint) and for adverse events (except the subgroups based on MDS-UPDRS Motor scores). Selected baseline data will be presented for the subgroups as well. Subgroups with too few patients per treatment arm (e.g. less than 4) may not be analysed.

- Patients with the baseline MDS-UPDRS MOTOR score (at the last titration visit at which the randomized dose is given up through TV6) less than or equal to the median versus the patients with the baseline score above the median
- Patients with the change from pre-dose to 30 minutes post-dose in MDS-UPDRS MOTOR score at baseline less than or equal to the median versus the patients with the change score above the median
- Non-elderly (< 65 years) versus elderly (≥ 65 years) patients

- Male versus female patients
- Race (any race category with less than 4 patients will be combined into an "Other" category for analysis) patient
- Dose level required to achieve a full "ON" response (10, 15, 20, 25, 30 or 35 mg)
    - For the purpose of efficacy analyses, this dose level is defined as the randomized dose level
    - For the purpose of the safety analysis using the Safety population, this dose level is defined as the highest dose level of APL- 130277 received during the Dose Titration Phase or Maintenance Treatment Phase.
    - For the purpose of the safety analysis using the Maintenance Phase Safety population, this dose level is defined as the highest dose level of study medication (APL- 130277 or placebo) received during the Maintenance Treatment Phase.

The following subset will be analyzed separately for demographic and disease history characteristics, and key safety outcomes during the dose titration phase:

- Enrichment failures defined as patients in the Safety population who were not randomized into the maintenance phase.

In addition, the following subsets will be analyzed separately for demographic characteristics and primary and key secondary efficacy endpoints:

- The subset excluding patients with a high baseline MDS-UPDRS MOTOR pre-dose assessment score, defined as a score > 55 points
- The subset including patients who were randomized to dose levels of 30 mg or less will be analyzed separately for demographic characteristics and primary and key secondary efficacy endpoints.

10 Disposition

The patient disposition will be summarized as follows and presented for each treatment group, as applicable, and overall. The percentages will be calculated based on the number of enrolled patients (i.e. patients who have signed the informed consent), unless otherwise specified.

- The number of patients screened (i.e. the number of patients in the AAP)
- The number (%) of patients who failed screening (% calculated from the AAP), including the distribution of reasons for failing the screening. If a patient failed screening multiple times, then all reasons for screen failures will be displayed. Percentages for these reasons will be calculated based on the total number of screen failure events
- The number (%) of patients enrolled into the study (% calculated from the AAP)
- The number (%) of patients who received at least one dose of study medication (Safety Population)

- The number (%) of patients who discontinued the study prematurely before Randomization, presented by highest dose level received during the Dose Titration Phase (10 mg, 15 mg, 20 mg, 25 mg, 30 mg or 35 mg and total), including the distribution of reasons for discontinuations
- The number (%) of patients randomized to the Maintenance Treatment Phase (ITT population)
- The patients who were randomized to the Maintenance Treatment Phase will be presented as a cross-tabulation of the highest dose level received during the Dose Titration Phase and the randomized dose level (10 mg, 15 mg, 20 mg, 25 mg, 30 mg or 35 mg) (% calculated from the ITT Population)
- The number (%) of patients who received at least one dose of study medication during the Maintenance Treatment Phase, i.e. Maintenance Phase Safety Population (% calculated from the ITT Population)
- The number (%) of patients in the mITT Population (% calculated from ITT Population)
- The number (%) of patients in the Completer Population (% calculated from ITT Population)
- The number (%) of patients in the Per Protocol Population (% calculated from ITT Population)
- The number (%) of patients who completed the study (% calculated from ITT Population)
- The number (%) of patients who discontinued the study prematurely after Randomization including the distribution of reasons for premature discontinuations after Randomization (% calculated from ITT Population)

Patient completion will also be summarized by visit and presented for each treatment group.

11 Demographic and other baseline characteristics

Demographics and screening/baseline characteristics will be summarized descriptively for the mITT, maintenance phase safety and Safety populations, by treatment group and/or overall. The following variables will be summarized:

- Demographics: (age (continuous), age categorized as <65 years versus ≥65 years, gender, ethnicity, race, height, weight, BMI, country)
- Smoking history
- Cognitive status: Mini–Mental State Examination (MMSE) total score (as categorical, % of patients with a score of 30, 29, 28, 27, 26 or <26)
- Modified Hoehn and Yahr scale in “ON” state
- MDS-UPDRS Total Score: Sum of MDS-UPDRS Part I (Non-Motor Aspects of Experiences of Daily Living), Part II (Motor Aspects of Experiences of Daily Living) and Part III (Motor Examination) assessed in an “OFF” state prior to L-Dopa administration at SV
- Baseline (SV) MDS-UPDRS Part I Score

- Baseline (SV) MDS-UPDRS Part II Score
- MDS-UPDRS Part III Score assessed in an "OFF" state prior to L-Dopa administration at SV
- MDS-UPDRS part III (Motor Examination) score at baseline pre-dose
- MDS-UPDRS part III (Motor Examination) score at baseline: 15, 30, 45, 60, 90 minutes post-dose
- Change in MDS-UPDRS part III (Motor Examination) score at baseline from pre-dose to: 15, 30, 45, 60, and 90 minutes post-dose after APL-130277 administration
- Percentage of patients with a patient-rated full "ON" response within 30 minutes at baseline after APL-130277 administration
- Total daily L-Dopa dose at baseline (PD medications reported by the patient as ongoing at first dose in titration phase will be added to calculate total daily dose)

Demographic characteristics and other baseline data will also be summarized by subgroups as defined in section 9.5 of this document

Baseline is defined as the last observed value before the first dose in titration phase, unless specified otherwise.

12 Concomitant medication/treatment

All medications will be classified using the Anatomical Therapeutic Chemical (ATC) classification codes and preferred drug names from the World Health Organization Drug Dictionary (WHO-DD).

Medications with a stop date before the first date of study drug dosing will be considered prior medications. Medications with start date or stop date on or after the first date of study drug dosing will be considered concomitant medications. The prior and concomitant medications will be summarized separately.

Summaries of prior and concomitant PD treatment medications (medications which start with ATC code N04) will be presented in tabular form using the ATC Level 4 and preferred term. Other prior and concomitant medications will be presented in tabular form using the ATC Level 1, ATC Level 2, and Preferred Term (PT). Frequencies and percentages of patients receiving medications will be presented by treatment group and overall. The tables will be sorted by overall descending frequency of ATC Level(s) and then, within an ATC Level, by overall descending frequency of PT.

If the medication start date is incomplete, then it will be imputed as follows for the purpose of determining concomitant use:

- If the start date is completely missing, the start date will be equal to the first dose date. However, if the stop date is not missing and is before the first dose date, then the stop date will be used instead.

- If the start day is missing: Check if month is same as month of first dose date of study treatment. If yes, impute to first dose date of study treatment; else impute First day of the month.

- However, if the stop date is not missing and is before the date of the first dose of study treatment or the imputed start date, then the stop date will be used instead. If the start day and month are missing: Check if year is same as the year of the first dose date of study treatment. If yes, impute to first dose date of study treatment; else impute to first day of the first month (January).

However, if the stop date is not missing and is before the date of the first dose of study treatment or the imputed start date, then the stop date will be used instead. If the medication stop date is partial, then it will be imputed as follows for the purpose of determining concomitant use:

- If the stop date is completely missing and the medication is not ongoing, the stop date will be equal to the last dose date or date of completion/withdrawal, whichever is the latest.
- If the stop day is missing, the last day of the month will be used. If resulted imputed stop date is after the latest of last dose date or date of completion/withdrawal, then the latest of last dose date or date of completion/withdrawal will be used.
- If the stop day and month are missing, then the last day of the last month (December) will be used. If resulted imputed stop date is after the latest of last dose date or date of completion/withdrawal, then the latest of last dose date or date of completion/withdrawal will be used.

In addition, the total daily levodopa dose (mg) will be summarized at baseline (PD medications reported by the patient as ongoing at first dose in titration phase will be added to calculate total daily dose) with descriptive statistics.

The summary of concomitant medications will be done for the mITT and Safety population.

13 Medical history

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The medical history data will be summarized with frequencies and percentages of patients with at least one medical history item, and patient frequencies and percentages on the System Organ Class (SOC) and PT levels. The number of events will also be summarized. The table will be sorted by overall descending frequency of SOC and then, within a SOC, by overall descending frequency of PT.

PD history (time since diagnosis of PD measured in years at time of first dose at TV1, presence of a rest tremor at the time of diagnosis, time since onset of motor fluctuations, type of OFF episodes experienced, number of OFF episodes/day, typical length of OFF episodes)

The summary of medical and PD history data will be done for the mITT and Safety population.

14 Analysis of efficacy

14.1 Primary endpoint

The primary endpoint of the study is the mean change from pre-dose in MDS-UPDRS MOTOR score after 30 minutes at 12 weeks (MV4). The difference between APL-130277 and placebo at MV4 will be estimated using a MMRM. The model will include the observed change from pre-dose MDS-UPDRS MOTOR score values after 30 minutes at MV1, MV2, MV3 and MV4 as the response values (i.e., no imputation will be done). The treatment difference at 12 weeks will be estimated using contrasts. The MMRM model will include the treatment group (APL-130277 or placebo), visit (MV1, MV2, MV3 and MV4) and the interaction between the treatment group and visit as fixed factors. The change from pre-dose in MDS-UPDRS MOTOR score after 30 minutes at the last titration visit at which the randomized dose is given up through TV6 will be used as a baseline covariate in the model.

An unstructured covariance structure will be applied for MMRM. In case the model will not converge with the unstructured covariance structure, the heterogeneous Toeplitz structure (TOEPH) will be used instead. In case the model will not converge with the heterogeneous Toeplitz structure, heterogeneous compound symmetry (CSH) will be used instead. The denominator degrees of freedom will be computed using the Kenward-Roger method.

The least square (LS) mean, standard error, and LS mean difference between APL-130277 and placebo group at MV4 along with the 95% confidence interval (CI) will be provided. The P-value for the hypothesis testing will also be provided. Treatment difference will be assessed with a 2-sided alpha level of 0.05, unless specified otherwise. The SAS code planned for the analysis is outlined below.

```
proc mixed data=&data;
class trtp avisit usubjid;
model chgpre=basepre trtp avisit trtp*avisit / ddfm=kr;
repeated avisit / subject=usubjid(trtp) type=un;
lsmeans trtp*avisit / cl;
run;
```

The mITT population will be used for the primary efficacy analysis. Observed and change from pre-dose MDS-UPDRS MOTOR scores will be summarized by visit and time point for the mITT population as well.

Subgroup analyses will be performed for factors defined in section 9.5 of this document. Each subgroup will be analyzed separately using the same methods as in the primary analysis. In addition, for each of the subgroup factors, a MMRM model similar to the primary model will be used, including additional fixed factors for the subgroup variable and the interaction between the treatment group and subgroup variable. The influence of each subgroup factor will be investigated using the p-value for the interaction term calculated with this model.

14.1.1 Patient Global Sensitivity analyses of the primary endpoint

The following sensitivity analyses will be performed for the primary endpoint.

- Completer analysis: The analysis as specified above will be repeated for the Completer Population.
- Per Protocol analysis: The analysis as specified above will be repeated for the Per Protocol Population.
- MI analysis with Missing At Random assumption: MI techniques based on Pattern Mixture Models (PMM) will be applied as a sensitivity analysis in the mITT population. This methodology will structure data based on missing data patterns. The method will be based on a missingness pattern having a monotone structure, i.e. if among the observations over time one data value is missing, all other values after this missing value will also be treated as missing. For patients with intermittent missing values, before performing MI based on the PMM, it will be necessary to create a monotone missingness pattern. Intermittent missing values will be imputed using the Markov Chain Monte Carlo (MCMC) methodology which assumes a multivariate normal distribution over all variables included in the imputation model. The MI procedure in SAS will be used for this purpose and this first MI step is planned to be repeated 100 times, creating several different datasets with a monotone missing data structure. Seed value of 201507 will be used in the MI procedure. The imputation is based on the missing at random (MAR) assumption, i.e. the missing data are assumed to follow the same model as the other patients in their respective treatment arm.

  After this, the remaining missing data can be imputed using a method for monotone missingness, also based on the MAR assumption. Thus, for each of the created datasets with a monotone missing data pattern, the MI procedure in SAS will be used to impute missing values based on a sequential procedure reflecting the monotone missing data pattern. Patients with the first missing value occurring at MV1 will have their missing MV1 value replaced by an imputed value from a regression model with treatment group and the change from pre-dose MDS-UPDRS MOTOR score at baseline as explanatory variables. In the next step, patients with their MV2 value missing will have their missing MV2 value replaced by an imputed value from a regression model with treatment group, the change from pre-dose MDS-UPDRS MOTOR score at baseline and the MV1 value as explanatory variables. Similar procedure will be used to replace the missing values at MV3 and MV4.

  The imputed datasets generated with the approach described above do contain only non-missing values and are used as input in the model for the sensitivity analysis of the primary endpoint. MMRM models similar as described above will thus be run on each of the generated imputed datasets and the difference between the treatment groups at MV4 will be estimated. The MMRM model will be similar to the primary analysis. Finally, the MIANALYZE procedure in SAS will be applied to combine the results from these several datasets to derive an overall estimate of the treatment difference at MV4. In addition to the estimates, corresponding 95% confidence intervals and p-values will be calculated.

- MI analysis with Missing Not At Random assumption (Placebo group based imputation): Another MI analysis will be performed with the assumption that the data is not missing at random. Placebo group based assumption will be used, i.e. the trajectories of the patients are assumed to follow the placebo group after the discontinuation. Methods similar to the procedure described above will be used in the mITT population. However, the missing values will be imputed using Placebo based imputation.

- MI analysis with Missing Not At Random assumption (Tipping point based imputation): If the primary analysis significantly favors APL-130277, another MI analysis will be performed in the mITT population with the assumption that the data in the APL-130277 group is not missing at random. A tipping point based assumption will be used, i.e. the trajectories of the patients in the APL-130277 group after withdrawal are assumed to be worse by an amount of delta. After the MI using the MAR assumption, as defined above has been done, the amount of delta will be added to each imputed value in the APL-130277 group. Successively harsher deltas will be imposed on the imputed values in the APL-130277 group, starting with a UPDRS increment (worsening) of 0.5 points. The delta is further increased in the steps of 0.5 points (1.0, 1.5, 2.0, …) until the statistical significance is lost, i.e. until the p-value becomes >0.05. For the placebo group, the MI using MAR assumption will be used.

- Comparability of the pre-dose values: The primary endpoint is defined as the change from the pre-dose value of the corresponding day. This definition is based on the assumption that the study treatment does not influence the pre-dose values during the study. This assumption will be investigated by tabulating the pre-dose values by visit and treatment group with descriptive statistics in the mITT population.

- Last observation carried forward (LOCF): In this analysis, the missing values will be replaced by the previous visit change values at the 30 min post-dose time point carried forward. The LOCF-imputed values at each visit will be compared using an Analysis of Covariance (ANCOVA) model with treatment group as a fixed factor. The change from pre-dose in MDS-UPDRS MOTOR score after 30 minutes at the last titration visit at which the randomized dose is given up through TV6 will be used as a covariate in the model. This analysis will be performed in the mITT population. Similarly, the pre-dose values corresponding to the change values will be carried forward for the purpose of summarizing the data with descriptive statistics.

- Responder analysis based on the MDS-UPDRS MOTOR scores. Response is defined as an improvement of at least 30% decrease in MDS-UPDRS MOTOR score from the pre-dose value at 15, 30, 45, 60 and 90 minutes. The number and proportion of responders at each time point and the cumulative number and proportion of patients having responded at least once by each time point will be tabulated with descriptive statistics at each visit.

14.1.2 Graphical summaries of the primary endpoint

The following graphs will be generated using the estimates calculated with the MMRM:

- Estimated mean changes (with SEM) from pre-dose to 15, 30, 45, 60 and 90 minutes by treatment group. Separate graph will be produced for each visit (MV1, MV2, MV3 and MV4). The x-axis includes the time of the assessment (15, 30, 45, 60 or 90 minutes) and the y-axis the change from pre-dose. LS means estimates and SEMs of these estimates from the primary MMRM will be used for the display.

- Estimated mean changes (with SEM) from pre-dose to 30 minutes by treatment group and visit. The x-axis includes the visit (MV1, MV2, MV3 or MV4) and the y-axis the change from pre-dose.

14.2 Key Secondary endpoint

The key secondary endpoint, percentage of patients with a patient-rated full "ON" response within 30 minutes at MV4 will be analyzed in the mITT population using the GLIMMIX procedure for binomial data with logit link. This analysis will use the observed values from MV1, MV2, MV3 and MV4 without any imputation as the response. The model will include the treatment group (APL-130277 or placebo), visit (MV1, MV2, MV3 and MV4) and the interaction between the treatment group and visit as fixed factors. The "ON/OFF" assessment at the last titration visit at which the randomized dose is given up through TV6 will be used as a covariate. An unstructured covariance structure will be used for the repeated measures. In case the model will not converge with the unstructured covariance structure, the heterogeneous Toeplitz structure (TOEPH) will be used instead. In case the model will not converge with the heterogeneous Toeplitz structure, heterogeneous compound symmetry (CSH) will be used instead. The odds ratio (APL-130277 vs. Placebo), 95% confidence intervals for the odds ratio and p-value will be provided.

The sas code planned for this analysis is

```
Proc glimmix data=&data;
class treatment visit subject baseline;
model response=treatment visit treatment*visit baseline /dist=binary link=logit ddfm=kr;
random visit/subject=subject type=un residual;  *if the model does not converge, try a different type by the following order: type=UN, type=TOEPH, type=CSH*/
lsmeans treatment*visit/cl ilink diffs oddsratio;
run;
```

As a sensitivity analysis, a Cochran-Mantel-Haenszel test stratified by the "ON/OFF" assessment within 30 minutes at the last titration visit at which randomized dose given. In case of missing data at MV4, the patient is considered as not reaching the full "ON". The percentage of patients with a patient-rated full "ON" response at each visit (MV1, MV2, MV3, MV4) and time point (15, 30, 45, 60 and 90 minutes) will also be summarized descriptively by treatment group. The analysis will be conducted in the mITT population.

14.2.1 Graphical summaries of percent of patients with patient-rated full "ON"

The following graphs will be generated for the percentage of patients with a patient-rated full "ON" response, as defined above:

- Percentage of patients with a patient-rated full "ON" response at 15, 30, 45, 60 and 90 minutes by treatment group. Separate graph will be produced for each visit (MV1, MV2, MV3 and MV4). The x-axis includes the time of the assessment (15, 30, 45, 60 or 90 minutes) and the y-axis the proportion of patients.

- Cumulative number of patients with a patient-rated full "ON" response within 15, 30, 45, 60 and 90 minutes by treatment group. Separate graph will be produced for each visit (MV1, MV2, MV3 and MV4). The x-axis includes the time of the assessment (15, 30, 45, 60 or 90 minutes) and the y-axis the cumulative proportion of patients reaching the response at least once by the time point.

- Percentage of patients with a patient-rated full "ON" response at 30 minutes by treatment group. The x-axis includes the visit (MV1, MV2, MV3 or MV4) and the y-axis the proportion of patients.

14.3 Other secondary endpoints

All the secondary endpoints defined below will be analyzed in mITT population.

1. The percentage of instances where a full "ON" response was achieved at 30 minutes after self-administration of study medication based on the home dosing diary entries.

   The percentages calculated separately for each patient for each visit (2 days prior to each visit) for the period between MV2 and MV4 (1 value per subject per visit) will be used as response variables in the statistical analysis. The percentage values will be compared between APL-130277 and placebo using a MMRM similar to the one used for the primary endpoint.

2. Mean change from pre-dose in MDS-UPDRS MOTOR score at 15, 45, 60 and 90 minutes at the 12 week visit (MV4) of the Maintenance Treatment Phase.

   These endpoints will be analyzed using a model similar to the MMRM used for the primary endpoint with corresponding timepoint for the covariate.

3. Time (in minutes) to when study medication is starting to have an effect at the 12 week visit (MV4) of the Maintenance Treatment Phase.

   This variable will be analyzed as a time-to-event endpoint. The time to effect at MV4 will be described using the Kaplan-Meier method, including the estimate of the median time to effect and corresponding 95% confidence intervals. In case the median time cannot be estimated in both treatment groups, the time to 25% of the patients having an effect will be estimated additionally. Furthermore, the Kaplan-Meier estimates at 5, 10, 15, … minutes will be summarized by 5 minute increments.. However, time points beyond 60 minutes will not be summarized. Cox proportional hazards model with treatment group as a factor will be used to compare APL-130277 and placebo.

The difference will be estimated with a hazard ratio along with a 95% CI for the hazard ratio. For the patients with missing data at MV4, the last available observation from visits MV1, MV2 or MV3 will be used instead. If all visits, MV1 – MV4, are missing, the patient will be considered censored at 90 minutes for this analysis.

4. Percent of patients with a patient-rated full "ON" response within 30 minutes, whose duration from time when study medication begins to have an effect until their "OFF" (if applicable) lasts for at least 30 minutes at the 12 week visit (MV4) of the Maintenance Treatment Phase.

   This endpoint will be analyzed using similar methods as used for the key secondary endpoint including the sensitivity analysis.

5. Clinical Global Impression of Improvement (CGI-I) post dosing.

   The proportions of patients who improved (defined as very much improved, much improved or minimally improved) will be tabulated by treatment group and visit. This summary will be complemented by the distribution of each response category (very much improved, much improved, minimally improved, no change, minimally worse, much worse, very much worse) tabulated by the treatment group and visit based on the observed results. The difference between APL-130277 and placebo at MV4 in patients who improved versus did not improve will be analyzed using a Cochran-Mantel-Haenszel test stratified by the CGI-S assessment at MV1. In case the distribution of the CGI-S assessment at MV1 is skewed, categories will be combined (e.g. category 1-3 vs 4 vs 5-7).

6. Patient Global Impression of Improvement (PGI-I) post dosing.

   This endpoint will be analyzed with methods similar to those used for CGI-I and proportion of patients who improved will be defined in a similar way. The Cochran-Mantel-Haenszel will be stratified by the PGI-S assessment at MV1. The PGI-S assessment at MV1 will be used to stratify the Cochran-Mantel-Haenszel test and categorized, if needed using approach similar to CGI-S.

7. Change from baseline in Parkinson's Disease Questionnaire-39 (PDQ-39).

   The PDQ-39 summary index will be analyzed with methods similar to those used for the primary endpoint with the exception of using the PDQ-39 score at SV as the baseline covariate. In addition, the sub-scores will be summarized with descriptive statistics.

8. Mean change from baseline in MDS-UPDRS – Part II: Motor Aspects of Experiences of Daily Living.

   This endpoint will be analyzed at MV4 with an ANCOVA model with treatment group as a fixed factor and the baseline (Screening Visit) value as a covariate.

   In addition to the part II score, each categorical question of MDS-UPDRS – Part IV (q4.1 Time Spent with Dyskinesias, q4.2 Functional Impact of Dyskinesias, q4.3 Time Spent in the OFF State, q4.4 Functional Impact of Fluctuations, q4.5 Complexity of Motor Fluctuations, q4.6 Painful OFF State Dystonia; each questions scored from 0 to 4) will be tabulated separately, question by question. Both the distribution of categories and the categorical change from baseline (Screening

Visit), categorized as improved, no change or worsened will be tabulated. In addition, the % of Dyskinesia time, % of OFF time and % of Dystonia time collected on the Part IV questionnaire will be summarized with descriptive statistics, both as absolute values and as changes from baseline.

14.4 Other patient-reported endpoints

1. Change in sleep measures on the Epworth Sleepiness Scale (ESS).

   The changes from SV to MV1, MV2, MV3, and MV4 in the total ESS score will be summarized with descriptive statistics. In addition, the proportion of patients who are sleepy (score of 10 or more) or very sleepy (score of 18 or more) will be tabulated.

2. Change in European Quality of Life – 5 Dimensions (EQ-5D).

   The distribution of responses for the health state dimensions will be tabulated. The changes from SV to MV1, MV2, MV3, and MV4 in health state thermometer scale and health state index score will be tabulated with descriptive statistics.

15 Analysis of safety and tolerability

The population used for safety analyses will be the Safety Population and Maintenance Phase Safety Population. In general, the safety data will be presented separately for the Dose Titration Phase (where applicable, using the Safety Population) and for the Maintenance Treatment Phase (using the Maintenance Phase Safety Population). The safety analysis of the Dose Titration Phase will summarize the safety data for the total Safety population while the analysis of the Maintenance Treatment Phase will focus on the comparison of the APL-130277 treated patients versus placebo. In addition, separate summaries will be prepared for all patients receiving APL-130277 either during the Dose Titration Phase or Maintenance Treatment Phase by pooling the data from the two periods during the APL-130277 exposure (using the Safety Population).

15.1 Extent of exposure

The following information will be summarized by treatment group (when applicable) and overall. In addition, these summaries will be broken down by the dose level of APL-130277 (10 mg, 15 mg, 20 mg, 25 mg, 30 mg or 35 mg; see Section 9.5).

- The number of patients exposed to study treatment, defined separately for the Dose Titration Phase (number of single dose exposures will be derived) and for the Maintenance Treatment Phase (repeated dose exposure)
- The following summaries will be done only for the Maintenance Treatment Phase (repeated dose exposure)

- Duration of exposure (days): Exposure to APL-130277 and Placebo in Maintenance Phase = date of last dose received in Maintenance Phase – date of first dose received in Maintenance Phase + 1
- Total exposure to study treatment, expressed as person years (sum of duration of exposure to study treatment over all patients in days divided by 365.25, classified by treatment group)
- Proportion of doses taken at days with 5, 4, 3, 2, 1 doses taken and average daily dose (during the days when the information was collected)
- Proportion of patients using >5, 5, 4, 3, 2, 1 or 0 doses/day at least once (during the days when the information was collected).

15.2 Compliance

Treatment compliance will be assessed by summarizing the percentage of patients taking more than 5 doses of study medication per day at least once during the study. The data will be summarized by treatment group during the Maintenance Treatment Phase.

15.3 Adverse events

All AEs will be coded using MedDRA Version 19.1.. Treatment-emergent adverse events (TEAEs) are defined as all AEs that start after the patient receives the first dose of study treatment. Dose Titration Phase TEAEs are defined as all AEs that start on or after the date of the first dose of APL-130277 during the Dose Titration Phase but before the date of the first dose of study medication (APL-130277 or placebo) during the Maintenance Treatment Phase. Maintenance Treatment Phase TEAEs are defined as all AEs that start on or after the first dose of study medication (APL-130277 or placebo) during the Maintenance Treatment Phase. Events will be classified as drug-related if the AE is classified as possibly, probably, or certainly related to study drug.

Separate summaries will be generated for TEAEs of special interest, tabulated by category and PT using the categories as specified below (Narrow terms will be used for all SMQs and all preferred terms under the specified HLT or HLGT as defined in the MedDRA dictionary will be included):

- Hypotension, orthostatic hypotension: defined as all TEAEs with HLGT “decreased and non-specific blood pressure disorders and shock”
- Syncope: defined as all TEAEs with any of the following PTs
  - “Hypotonic-hyporesponsive episode” (MedDRA code 10021121)
  - “Altered state of consciousness” (MedDRA code 10001854)
  - “Depressed level of consciousness” (MedDRA code 10012373)
  - “Hypokinesia” (MedDRA code 10021021)
  - “Hypokinesia neonatal” (MedDRA code 10021022)
  - “Hyporesponsive to stimuli” (MedDRA code 10071552)
  - “Loss of consciousness” (MedDRA code 10024855)
  - “Neurogenic shock” (MedDRA code 10058119)

- "Presyncope" (MedDRA code 10036653)
- "Shock" (MedDRA code 10040560)
- "Shock symptom" (MedDRA code 10040581)
- "Syncope" (MedDRA code 10042772)
- "Unresponsive to stimuli" (MedDRAcode 10045555)

Or HLGT of "Seizures (incl subtypes)"

- Falls & injuries: defined as all TEAEs meeting the criteria for the standardized MedDRA Query (SMQ) "Accidents and injuries"
- Dyskinesias: defined as all TEAEs meeting the criteria for the SMQ "Dyskinesia"
- Hallucinations and psychotic behaviors: defined as all TEAEs meeting the criteria for the SMQ "Psychosis & psychotic disorders"
- Impulse control disorders: defined as all TEAEs meeting the following criteria.
  - HLT "Impulse Control Disorders"
  - HLT "Paraphilia and paraphilic disorders" or "Sexual desire disorders"
  - Binge eating: PT "Binge Eating" (MedDRA code 10004716)
  - Gambling: PT "Gambling" (MedDRA code 10017655) or Gambling Disorder (MedDRA code 10078070)
  - Compulsive shopping: PT "Compulsive shopping" (MedDRA code 10067948) or any TEAE with verbatim term including the text "spending"
- Daytime sudden onset of sleep: defined as all TEAEs with HLGT "Sleep disorders and disturbances"
- QT prolongation and ventricular arrhythmias: defined as all TEAEs meeting the criteria for the SMQ "Torsade de pointes /QT prolongation"
- Acute Coronary Syndrome, Myocardial infarction, Angina: defined as all TEAEs meeting the criteria for the SMQ "Myocardial infarction"
- Suicidal ideation & attempts: defined as all TEAEs meeting the criteria for the SMQ "Suicide/self-injury"
- Melanoma: defined as all TEAEs meeting the criteria for the SMQ "Skin malignant tumors"
- Stomatitis, Oral ulcers, Oral irritation: defined as all TEAEs meeting the criteria for the SMQ "Oropharyngeal disorders"
- Allergic/sensitivity response to the formulation: defined as all TEAEs meeting the criteria for the SMQ "Hypersensitivity"

Events with a missing start time, but with a start date equal to the date of first dose of study treatment will be considered treatment-emergent. If the AE start date is incomplete, it will be imputed as follows for the purpose of determining TEAE:

- If the start date is completely missing, the start date will be equal to the date of the first dose of study treatment. However, if the stop date is not missing and is before the date of the first dose of study treatment, then the stop date will be used instead.
- If the start day is missing: Check if month is same as month of first dose date of study treatment. If yes, impute to first dose date of study treatment; else impute First day of the month.
- However, if the stop date is not missing and is before the date of the first dose of study treatment or the imputed start date, then the stop date will be used instead. If the start day and month are missing: Check if year is same as the year of the first dose date of study treatment. If yes, impute to first dose date of study treatment; else impute to first day of the first month (January).

However, if the stop date is not missing and is before the date of the first dose of study treatment or the imputed start date, then the stop date will be used instead.

The original date and time will be shown on all listings of AEs. Listings will be provided for all AEs, serious TEAEs, non-serious TEAEs, TEAEs leading to drug withdrawal, TEAEs leading to drug interruption, TEAEs leading to dose reduction, and AEs leading to deaths. The listings will display study day, calculated as the AE start date – date of first dose in the relevant study phase (Dose Titration or Maintenance Treatment) + 1 for events occurring on or after the first dose in the Dose Titration phase, and as AE start date – date of first dose in the Dose Titration phase for AEs occurring prior to the first dose in the Dose Titration Phase.

TEAEs will be summarized by SOC and PT and by treatment group when applicable. TEAEs with onset after the last dose of the study treatment period are attributed to the treatment received during the treatment period. Both event and patient counts, where applicable, will be summarized. The counts will be complemented by percentages calculated for the patient counts unless otherwise specified.

- An overall summary of the number and percentage of patients reporting TEAEs and the number of TEAE events, drug-related TEAEs, severe TEAEs, serious TEAEs, non-serious TEAEs, TEAEs leading to drug withdrawal, TEAEs leading to drug interruption, TEAEs leading to dose reduction, and TEAEs leading to death
- TEAEs by SOC and PT, both as event and patient counts
- TEAEs by PT, both as event and patient counts
- Drug-related TEAEs by SOC and PT, both as event and patient counts
- Severe TEAEs by SOC and PT, both as event and patient counts
- TEAEs by SOC, PT and severity, as event counts; percentages will be calculated for the event count out of total number of events
- TEAEs by SOC, PT and relationship, as event counts; percentages will be calculated for the event count out of total number of events
- Serious TEAEs by SOC and PT, both as event and patient counts

- TEAEs leading to drug withdrawal by SOC and PT, both as event and patient counts
- TEAEs leading to death by SOC and PT, both as event and patient counts
- TEAEs leading to drug interruption by SOC and PT, both as event and patient counts
- TEAEs leading to dose reduction by SOC and PT, both as event and patient counts
- TEAEs of special interest by category and PT, both as event and patient counts
- Non-serious TEAEs by SOC and PT, both as event and patient counts
- Non-serious TEAEs with >3% cumulative patient incidence in any arm/group by SOC and PT, both as event and patient counts

The tables will be sorted by overall descending frequency of SOC and then, within a SOC, by overall descending frequency of PT based on the patient count for the APL-130277 column. If only event count is presented, the sorting will be done based on the event count.

Furthermore, the adverse events will also be summarized by subgroups (age, gender, race, country, dose level) as defined in section 9.5 of this document. In addition, the following tables will be summarized by dose: TEAEs, Serious TEAEs, TEAEs leading to drug withdrawal, TEAEs leading to drug interruption, TEAEs leading to dose reduction, TEAEs leading to death, and TEAEs of special interest.

15.4 Laboratory safety variables

All clinical laboratory parameters will be converted to consistent units according to the International System of Units (SI) before summarization. The following will be summarized:

- Numeric laboratory parameters: Actual values and change from baseline (SV) to EOS, if applicable, for each parameter, will be summarized with descriptive statistics.
- Laboratory parameters which have an upper or lower reference range: Number and percentage of patients with low, normal or high (i.e., below, within or above reference range) values at each visit for each parameter will be summarized
    - These values will be presented as a shift table, i.e. the distribution of the three response categories at EOS visit will be classified by the baseline (SV) category.
- Categorical laboratory parameters: The distribution of the categories will be summarized by visit.

15.5 Vital signs

Vital sign measurements include heart rate (HR), respiratory rate (RR), blood pressure (BP) and body temperature. Vital signs will be measured at all scheduled study visits at various time points after the patient has been in a supine position for 5 minutes. BP will also be measured within three minutes of standing at all time points. In addition to the vital signs captured on the Case Report Form, the standing

minus supine values (standing minus supine systolic BP, standing minus supine diastolic BP) will be evaluated.

The following summaries will be done:

- Change from baseline (SV) to other visit pre-dose values for each parameter (including the standing minus supine values). The change from baseline to the End of Study visit values will also be presented.
- Actual values and change from pre-dose to post-dose, if applicable, at each visit for each parameter (including the standing minus supine values).

Orthostatic hypotension will be defined as a reduction in systolic BP of 20 mmHg or more, or a reduction in diastolic BP of 10 mmHg or more, for the standing measurement compared to the supine measurement. The proportion of patients with orthostatic hypotension will be tabulated by visit and time point.

15.6 ECG

A standard 12-lead ECG will be performed at all time points outlined in the protocol. ECGs will be performed in a semi-recumbent position and after 5 minutes of rest. The following parameters will be captured: Heart rate, PR interval, QRS interval, RR interval, QT interval, QTc Interval (Fridericia's correction) and QTc Interval (Bazett's correction). A triplicate 12-lead ECG will be performed at screening and the median value will be used for summary purposes.

The following summaries will be done:

- Change from baseline (SV) to other visit pre-dose values for each parameter. The change from baseline to the End of Study visit values will also be presented.
- Actual values and change from pre-dose to post-dose, if applicable, at each visit for each parameter.

The ECGs will be centrally read and deemed "Normal", "Abnormal, not clinically significant" and "Abnormal, clinically significant" and most severe assessment will be tabulated by visit and time point.

In addition, the QTc Intervals fulfilling the following criteria will be tabulated separately using Fridericia's correction and Bazett's correction:

- Values >500 msec
- Values increasing >15% from baseline if baseline value is ≥440 msec
- Values increasing >30% from baseline if baseline value is <440 msec
- Values increasing >30 msec from baseline
- Values increasing >60 msec from baseline
- At least one of the abnormalities listed above.

15.7 Other safety variables

The Columbia Suicide Severity Rating Scale (C-SSRS) is a measure of suicidal ideation and behavior. The rating scale has 4 general categories: suicidal ideation, intensity of ideation, suicidal behavior, and actual attempts. All C-SSRS data will be listed. The frequency and percentage of patients with each response for suicidal ideation, intensity of ideation, and suicidal behavior items will be summarized as appropriate by treatment group and visit.

The Questionnaire for Impulsive-Compulsive Disorders in Parkinson's disease – Rating Scale (QUIP-RS) is an instrument used to measure the extent of impulsive and compulsive behaviors in PD patients. The QUIP-RS consists of four questions which have to be answered for each disorder (gambling, sex, buying, eating, hobbyism, punding and PD medication use) on a 5-point Likert scale. Scoring range for each scale (i.e. disorder) is 0–16. The frequency and percentage of patients with positive response ("rarely" or higher) for each disorder will be summarized by treatment group and visit. Furthermore, the total Impulsive Control Disorder (ICD) score and the total QUIP-RS score will be summarized with descriptive statistics by treatment group and visit (both absolute values and changes from baseline).

The frequency and percentage of patients with physical examination findings will be summarized by treatment group, visit, time point (when applicable) and body system. In addition, the frequency and percentage of patients with each type of oropharyngeal cavity examination finding will be summarized by treatment group, visit, time point (when applicable) and location.

16 Other variables

Major protocol deviations will be tabulated including the frequency and percentage of patients with each type of deviation by treatment group. Deviations considered major will be identified as such prior to study unblinding. Summaries will be based on the Enrolled Population and provided for the Screening Phase and Dose Titration Phase. Summaries will also be presented for ITT population for the Maintenance Treatment Phase. All other data collected but not specifically mentioned including the Zarit Burden Interview will be listed.

17 Deviations from the analyses planned in the study protocol

This analysis plan presents the planned analyses in more detail and several details and additional endpoints have been added.

The definition of modified intention-to-treat population was changed from "All patients who are randomized, receive at least one post-randomization dose of study treatment (APL-130277 or placebo) and have efficacy data from at least one post randomization evaluation will comprise the mITT population' to "All patients who are randomized and receive at least one post-randomization dose of study treatment (APL-130277 or placebo) will comprise the modified Intention-To-Treat (mITT) population.

18 Execution of statistical analyses

Statistical analyses will be performed by INC Research/inVentiv Health and supervised by Sunovion and CLINTREX LLC.

19 Hardware and software

Statistical analysis, tables and patient data listings will be performed with SAS® version 9.3 or later for Windows (SAS Institute Inc., Cary, NC, USA).

20 References

Clinical Study Protocol: Version 2.0 (6 May, 2015), Company: Sunovion Pharmaceuticals Inc.