Statistical Analysis Plan

CTH-302
An Open-Label, Randomized, Crossover Trial utilizing a Single-Blinded Rater to evaluate APL-130277 compared to s.c. Apomorphine in Levodopa Responsive Subjects with Parkinson's Disease Complicated by Motor Fluctuations

Phase: Phase III
Author:
Version Number and Date: V0.4.01, 11 May 2021

Statistical Analysis Plan Signature Page

Statistical Analysis Plan V0.4.01 (Dated 11MAY2021) for Protocol CTH-302.

	Name	Signature	Date
Author:			
Position:			
Company:	Sunovion Pharmaceuticals Inc.		

Upon review of this document, the undersigned approves this version of the Statistical Analysis Plan, authorizing that the content is acceptable for the reporting of this study.

	Name	Signature	Date
Approved By:			
Position:			
Company:	Sunovion Pharmaceuticals Inc.		
	Name	***Signature***	
Approved By:			
Position:			
Company:	Sunovion Pharmaceuticals Inc.		

Modification History

Unique Identifier for this Version	Author	Significant Changes from Previous Authorized Version
1.0		Not Applicable – First Version
4.01		(1) Updated definition of the treatment Preference endpoint to be based on question 9a (Q9a) or 9b (Q9b) of the TPQ (see 15.2.1.2.). (2) Removal of blinded sample size re-assessment (b-SSR) due to enrollment challenges (including covid-19 pandemic).

Table of Contents

1. ABBREVIATIONS

AE	Adverse Event
ANCOVA	Analysis of Covariance
APL	APL-130277
AR	Autoregressive
ATC	Anatomical Therapeutic Chemical
AUC	Area Under the Curve
BP	Blood Pressure
CGI-S	Clinical Global Impression of Severity
CGI-I	Clinical Global Impression of Improvement
CI	Confidence Interval
C_{max}	maximum observed plasma concentration
C-SSRS	Columbia Suicide Severity Rating Scale
DMC	Data Monitoring Committee
ECG	Electrocardiogram
ENR	Enrolled population
EOS	End Of Study
EQ-5D	European Quality of Life – 5 Dimensions
HLGT	High-Level Group Term
HR	Heart Rate
IXRS	Interactive X Response System
KR	Kenward-Roger
L-Dopa	Levodopa
LOCF	Last Observation Carried Forward
LS	Least Square
MAR	Missing At Random
MCMC	Markov Chain Monte Carlo
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
MNAR	Missing not at random
OH	Orthostatic Hypotension
PD	Parkinson's Disease
PDQ-39	Parkinson's Disease Questionnaire-39
PE	Physical Examination
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Severity
PK	Pharmacokinetics
PT	Preferred Term
QUIP-RS	Questionnaire for Impulsive-Compulsive Disorders in Parkinson's disease – Rating Scale
RR	Respiratory Rate
SAE	Serious Adverse Event
SAF	Safety population
SAP	Statistical Analysis Plan
SAS	Statistical analysis system
sc/s.c.	Subcutaneous
SD	Standard Deviation
SEM	Standard Error of the Mean
SI	International System of Units
sl	Sublingual
SOC	System Organ Class
SV	Screening Visit
TPQ	Treatment Preference Questionnaire
TSQM	Treatment Satisfaction Questionnaire for Medication
TV	Titration Visit

MDS-UPDRS	Movement Disorder SocietyUnified Parkinson's disease Rating Scale
VAS	Visual Analog Scale
WHO-DD	World Health Organization Drug Dictionary

2. INTRODUCTION

This document describes the rules and conventions to be used in the presentation and analysis of efficacy and safety data for study CTH-302. It describes the data to be summarized and analyzed, including specifics of the statistical analyses to be performed.

This statistical analysis plan (SAP) is based on protocol version 4.01, dated 28 February 2020.

3. STUDY OBJECTIVES

3.1. PRIMARY OBJECTIVE

The primary objective is to demonstrate the efficacy of sublingual (sl) APL-130277 compared to subcutaneous (sc) apomorphine as a treatment of "OFF" episodes in subjects with Parkinson's Disease (PD) as measured by the change from pre-dose to 90 minutes post-dose in Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Part III score.

3.2. SECONDARY OBJECTIVES

The secondary objectives are to demonstrate the efficacy of sublingual (sl) APL-130277 compared to subcutaneous (sc) apomorphine as a treatment of "OFF" episodes in subjects with Parkinson's Disease (PD) as measured by:

- Durability of effect, defined as Investigator confirmed full "ON" within 30 minutes post-dose and at 90 minutes post-dose;
- Subject preference for APL-130277 after the subject has completed both APL-130277 and s.c. apomorphine treatment regimens;
- Subject confirmed durability of effect, defined as subject confirmed full "ON" within 30 minutes post-dose and at 90 minutes post-dose;
- Patient Global Impression of Change of "OFF" episodes (PGI-C).

3.3. PHARMACOKINETICS OBJECTIVES

To assess the pharmacokinetics (PK) for apomorphine and metabolites (apomorphine sulfate, norapomorphine, and others as deemed necessary) from plasma samples following dosing with sl APL-130277 and s.c. apomorphine.

3.4. SAFETY OBJECTIVES

To evaluate the safety and tolerability of sl APL-130277 at doses of 10 to 30 mg compared to s.c. apomorphine at doses of 2 to 6 mg during titration and the maintenance treatment period through the

assessment of: adverse events (AE) with attention to nausea, vomiting, hypotension, syncope, dyskinesia and impulse control disorders; physical examinations (PE) including oropharyngeal and injection site reactions; electrocardiogram (ECG) derived morphology and conduction parameters; vital signs including orthostatic hypotension (OH); frequency of domperidone use; clinically significant changes in laboratory tests; incidence of suicidal thoughts and actions as assessed by the Columbia Suicide Severity Rating Scale (C-SSRS); and impulsivity as measured by the Questionnaire for Impulsive Compulsive Disorders in Parkinson's Disease – Rating Scale (QUIP-RS).

3.5. STUDY ENDPOINTS

3.5.1. PRIMARY EFFICACY ENDPOINT

The primary efficacy variable is the change from pre-dose to 90 minutes post-dose in MDS-UPDRS Part III score after 4 weeks of dosing in each crossover period (assessed by the blinded-rater in-clinic at V3 and V6 of PART B).

3.5.2. SECONDARY EFFICACY ENDPOINTS

- Durability of effect, defined as an Investigator confirmed full "ON" within 30 minutes post-dose and at 90 minutes post-dose, after 4 weeks of dosing in each crossover period (assessed by the blinded-rater in-clinic at V3 and V6 of PART B).
- Subject preference for APL-130277 treatment as recorded for question 9 (Question 9a or 9b) of the TPQ. This assessment is scheduled to be performed after the subject has completed both APL-130277 and s.c. apomorphine treatment regimens (assessed in clinic at V6 of PART B).
- Subject confirmed durability of effect, defined as subject confirmed full "ON" within 30 minutes post-dose and at 90 minutes post-dose, after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B).
- Patient Global Impression of Change (PGI-C): Subject improvement of "OFF" episodes, defined as very much better, much better or a little better after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B).

3.5.3. OTHER EFFICACY ENDPOINTS (PART B)

- Clinical Global Impression of Improvement (CGI-I): Subject improvement, defined as very much improved, much improved, or minimally improved after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B).

- Change from pre-dose to 15, 30, 60, 90, and 120 minutes post-dose in MDS-UPDRS Part III score (assessed in-clinic V1, V2, V4, and V5). Change from pre-dose to 15, 30, 60 and 120minutes post-dose in MDS-UPDRS Part III score (assessed in-clinic V3 and V6).
- Investigator confirmed full "ON" at the 15, 30, 60, 90, and 120 minutes post-dose time points (assessed in-clinic V1, V2, V3, V4, V5, and V6).
- Subject confirmed full "ON" at the 15, 30, 60, 90, and 120 minutes post-dose time points (assessed in-clinic V1, V2, V3, V4, V5, and V6).
- Time to full "ON" and Time to partial "ON" as determined by the subject and Investigator (assessed in-clinic at all visits in PART B).
- Per the Expanded Home Dosing Diary, percent of episodes with a subject-rated full "ON" within 30 and at 90 minutes post-dose based on the 3 consecutive days prior to V2, V3, V5, and V6.
- Level of satisfaction as assessed by the Treatment Satisfaction Questionnaire for Medication (TSQM) (assessed in-clinic at V3 and V6).
- MDS-UPDRS – Part I Score: change from Screening to V2, V3, V5, and V6.
- MDS-UPDRS – Part II Score: change from Screening to V2, V3, V5, and V6.
- MDS-UPDRS – Part IV Score: change from Screening to V2, V3, V5, and V6.
- Improvement in amount of troublesome dyskinesia in the last month per the in-clinic dyskinesia questionnaire; defined as less than usual amount of troublesome dyskinesia after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6).
- Percent of "ON" episodes without troublesome dyskinesia (per expanded home diary) based on the 3 consecutive days prior to V2, V3, V5, and V6.
- Change in the PDQ-39 Total Index Score, and Mobility and Activities of Daily Living subscale scores from pre-dose V1 to after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6).
- European Quality of Life – 5 Dimensions (EQ-5D-5L) (assessed in-clinic at V1, V3, and V6).
- Ease of Use Questionnaire (assessed in-clinic at V3 and V6).

3.5.4. SAFETY ENDPOINTS

Evaluation of safety and tolerability of APL-130277 compared to s.c. apomorphine as measured by AEs, including assessment of oropharyngeal AEs and injection-site related AEs, 12-lead ECGs, vital signs including OH, clinical laboratory tests, and C-SSRS and QUIP-RS assessments.

3.5.5. PHARMACOKINETIC ENDPOINTS

Pharmacokinetic concentration-time data for apomorphine and metabolites (apomorphine sulfate, norapomorphine, and others as deemed necessary) will be evaluated and PK parameters (including but not limited to C_{max}, t_{max}, AUC_t, parent-to-metabolite ratios of C_{max} and AUC_t) will be estimated by noncompartmental methods from plasma samples using actual elapsed time from dosing. Details and methodology concerning the analysis of these variables will be provided in a separate PK analysis plan. A separate and stand-alone PK report will be provided.

4. STUDY DESIGN

4.1. GENERAL DESCRIPTION

This is a two-part study: Part A and Part B.

PART A consists of an open label, crossover titration phase where eligible subjects will be randomized to 1 of 2 treatment sequences in a 1:1 ratio to (1) sl APL-130277 followed by sc apomorphine or (2) sc apomorphine followed by sl APL-130277. APL-130277 sublingual films will be provided in 5 strengths: 10 mg, 15 mg, 20 mg, 25 mg, and 30 mg. Subcutaneous injection apomorphine will be provided in 5 strengths: 2 mg, 3 mg, 4 mg, 5 mg, and 6 mg.

Subjects will undergo dose titration with the first study treatment (APL-130277 or sc apomorphine) to tolerance and effect, i.e., the tolerable dose that turns the subject from the practically defined "OFF" state to the full "ON" state as determined by both the Investigator and subject. The subject will then be crossed over to the other study treatment (APL-130277 or sc apomorphine) and similarly titrated to tolerance and effect. The doses of APL-130277 and sc apomorphine determined as optimal for each subject in PART A will be used during PART B.

The titration of APL-130277 will start in the clinic at 10 mg, and if the subject fails to convert from a practically defined "OFF" to a full "ON" within 30 minutes (as assessed by both the Investigator and subject), the subject will be instructed on how to self-administer increasing doses of APL-130277 and continue the titration process at home. A schematic showing the flow of APL-130277 titration is shown in Figure1 of the protocol.

The site will contact the subject daily to determine titration status and query for safety issues. If the Investigator determines there are tolerability issues with APL-130277 in the home titration setting, the subject will be instructed to withhold further administration of APL-130277 and come to the clinic for an unscheduled clinic visit. Uptitration in 5 mg increment should continue at home until the subject achieves the optimal "ON" response. After the subject determines an optimal dose of APL-130277 that converts him/her from a practically defined "OFF" to a full "ON", the subject will return to the clinic to confirm the efficacy and tolerability of the selected dose.

Titration with sc apomorphine will take place in the clinic only, starting at 2 mg.

Following washout period of at least 1 day, eligible subjects will enter PART B of the study and be randomized to 1 of 2 treatment sequences in a 1:1 ratio.

Subjects will be instructed to continue with their regular PD medication regimen(s), and to dose themselves with their randomized open-label study treatment (APL-130277 or sc apomorphine) if they experience an "OFF" episode (e.g., morning akinesia, wearing "OFF", dose failure, sudden "OFF", etc.) during the waking day. Subjects will be instructed to dose up to 5 "OFF" episodes per day during the waking period and to complete a daily home dosing diary. Study drug doses (APL-130277 or sc apomorphine) must be separated by at least 2 hours.

During the 4-week open-label treatment period, subjects will return to the clinic for safety and efficacy assessments at 2-week intervals. In PART B, the MDS-UPDRS, CGI-S, CGI-I and ON/OFF assessment will be assessed by a rater blinded to the subject's study treatment.

The subject will then be crossed over to the other study treatment (APL-130277 or sc apomorphine) for an additional 4-weeks of open-label treatment, following the same assessment schedule.

A study schematic is presented in Figure 3 of the protocol. Details of the study assessments and other procedures to be performed at each visit are presented in Table 2 (Part A) and Table 3 (Part B), Schedule of Assessments, and Section 11, Study Assessments of the protocol. If necessary, subjects may return to the clinic at any time for an unscheduled visit.

4.2. METHOD OF ASSIGNING SUBJECTS TO TREATMENT GROUPS

The treatment schedule was generated by an independent biostatistician. Once a subject is deemed eligible to be randomized, the interactive X response system (IXRS) will perform the treatment assignment. Blocked randomization will be used with no stratification factors.

In PART A, subjects will be randomized to one of the following treatment sequences in a 1:1 ratio:

- Sequence 1: APL-130277 followed by sc apomorphine;
- Sequence 2: sc apomorphine followed by APL-130277.

In PART B, subjects will be randomized to one of the following treatment sequences in a 1:1 ratio:

- Sequence 3: APL-130277 followed by sc apomorphine;
- Sequence 4: sc apomorphine followed by APL-130277.

4.3. BLINDING

This is an open-label trial, however, in PART B, the MDS-UPDRS, CGI-S, CGI-I, and ON/OFF assessment will be assessed by a rater blinded to the subject's study treatment.

4.4. DETERMINATION OF SAMPLE SIZE

The sample size calculation is based on the primary endpoint of the study, the mean change from pre-dose in MDS-UPDRS Part III Motor Examination score at 90 minutes post-dose, evaluated at V3 and V6 for each treatment in PART B. Assuming that the discontinuation rate is 25% in PART A, a total of 106 subjects is planned to be randomized into PART A, so that at least 80 subjects are randomized into PART B. Assuming a 30% discontinuation rate in PART B, approximately 55 subjects are expected to complete PART B. With 55 subjects, the study has 90% power to detect a mean treatment difference between APL-130277 and sc apomorphine of 5.5 points for the change in MDS-UPDRS Part III score, assuming a standard deviation of 12 points for the period differences in PART B. The expected mean treatment difference of at least 5.5 points for the change in MDS-UPDRS Part III score was based on the results of the CTH-300 study.

4.5. CHANGES IN THE CONDUCT OF THE STUDY

The TPQ assessment of subject treatment preference was changed from being based on a visual analogue scale (VAS) to a 5-point Likert scale in protocol amendment 3.0. This change was made based on results of a preference validation study of the VAS performed by Sunovion which showed Parkinson's disease patients had difficulty completing the VAS compared to the new 5- point scale.

4.6. SCHEDULE OF EVENTS

Schedule of events can be found in Section 1 of the protocol (Table 2 for Part A and Table 3 for Part B).

4.7. CHANGES TO ANALYSIS FROM PROTOCOL

(1) Change to definition of subject treatment Preference endpoint to be based on question 9a (Q9a) or 9b (Q9b) of the TPQ (see 15.2.1.2.).

(2) Removal of blinded sample size re-assessment (b-SSR) due to enrollment challenges (including COVID-19 pandemic).

5. PLANNED ANALYSES

The following analyses will be performed for this study:

Final Analysis

5.1. DATA MONITORING COMMITTEE (DMC)

There will be no DMC for this study.

5.2. INTERIM ANALYSIS

No formal, interim analysis is planned.

5.3. FINAL ANALYSIS

All final planned analyses identified in this SAP will be performed by Syneos Biostatistics following Sponsor Authorization of this Statistical Analysis Plan, Database Lock, and Sponsor Authorization of Analysis Sets.

Some minor modifications may be necessary to the planned design of tables, figures, and listings to accommodate data collected during the actual study conduct.

The complete PK analyses to be performed will be specified and results reported in a separate document(s).

6. ANALYSIS POPULATIONS

The subjects included/excluded from each analysis set will be finalized at the data review meeting prior to the database lock, without regard to treatment.

6.1. ALL SUBJECTS ENROLLED [ENR] POPULATION

The all subjects enrolled (ENR) population will contain all subjects who provide informed consent for this study.

6.2. MODIFIED INTENT-TO-TREAT [MITT] POPULATION

The following modified Intent-to-Treat (mITT) populations will be defined:

- PART A mITT Population: The PART A mITT Population includes all subjects who were randomized and received at least one dose of either study drug in PART A. This population will be used for any efficacy analysis of PART A data according to randomized treatment.
- PART B mITT Population: The PART B mITT Population includes all subjects who were randomized and received at least one dose of either study drug in PART B. This population will be used for the efficacy analysis of PART B data according to randomized treatment.

6.3. SAFETY [SAF] POPULATION

The following safety populations will be defined:

- PART A Safety Population: The PART A Safety Population includes all subjects who received at least one dose of either study drug in PART A. The PART A Safety Population will be used for the safety analysis of PART A data according to treatment received.

- PART B Safety Population: The PART B Safety Population includes all subjects who received at least one dose of either study drug in PART B. The PART B Safety Population will be used for the safety analysis of PART B data according to treatment received.

6.4. OTHER ANALYSIS SET (S)

Not applicable.

6.5. PHARMACOKINETIC (PK) POPULATION

The PK population will be defined in the PK analysis plan.

7. GENERAL CONSIDERATIONS

All applicable data will be summarized separately for PARTs A and B, unless stated otherwise. Where appropriate, data will be summarized by visit week and/or time point in addition to treatment group. For PART B analyses by visit week, data from V1 and V4 will be labeled as Week 0; V2 and V5 will be labeled as Week 2; and V3 and V6 will be labeled as Week 4.

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 percentages. The total number of subjects in the treatment group overall (N) will be used as the denominator for percent calculations, unless stated otherwise. Significance testing will be 2-sided at the 0.05 level, unless otherwise specified.

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

Statistical analysis, tables and subject data listings will be generated using SAS® version 9.4 or later for Windows (SAS Institute Inc., Cary, NC, USA). Deviations from the statistical plan will be reported in the clinical study report, including the rationale for the change.

Reference Date and Study Day

Study Day will be calculated from the reference start date and will be used to show start day of assessments and events.

Reference start date is defined as the day of the first dose of study medication in each period of treatment, (Day 1 is the day of the first dose of study medication.

- If the date of the event is on or after the reference date then:
  Study Day = (date of event – reference date) + 1.
- If the date of the event is prior to the reference date then:
  Study Day = (date of event – reference date).

In the situation where the event date is partial or missing, Study Day, and any corresponding durations will appear missing in the listings.

Baseline

Unless otherwise specified, baseline is defined as the last non-missing measurement taken prior to reference start date (including unscheduled assessments). In the case where the last non-missing measurement and the reference start date coincide, that measurement will be considered pre-baseline, but Adverse Events (AEs) and medications commencing on the reference start date will be considered post-baseline.

Retests, Unscheduled Visits, and Early Termination

In general, for by-visit summaries, data recorded at the nominal visit will be presented. Unscheduled measurements will not be included in by-visit summaries but will contribute to the EOS/ EOT/ Endpoint/ LOCF value, or best/ worst case value where required (e.g. shift table).

In the case of a retest (same visit number assigned), the latest available measurement for that visit will be used for by-visit summaries.

If, for a particular treatment in Part A, the titration is repeated to re-confirm the dose (possibly due to an extended gap before enrolling in Part B), then the re-confirmation visits will be considered unscheduled and the by visit/dose table summaries will utilize the scheduled assessments associated with the planned treatment titration period.

Listings will include scheduled, unscheduled, retest and early discontinuation data.

8. STATISTICAL CONSIDERATIONS

8.1. ADJUSTMENTS FOR COVARIATES AND FACTORS TO BE INCLUDED IN ANALYSES

For the analysis of efficacy endpoints, the baseline or pre-dose value of the endpoint in question will generally be included as a model covariate, if available.

Specific use of covariates for statistical adjustment for each endpoint is discussed below where relevant.

8.2. MULTICENTER STUDIES

This study will be conducted at multiple centers in Europe utilizing a central randomization. Because of the large number of centers participating in this study, no by-center displays or adjustments are planned. Selected analyses by country/region may be performed, as described below.

8.3. MISSING DATA

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

- For the efficacy analysis, generally a likelihood-based modeling approach will be used to handle incomplete data using missing at random (MAR) assumption. The specific methods for handling missing data for each endpoint are specified in sections below.
- Sensitivity analysis for primary efficacy data will be conducted using a missing not at random (MNAR) approach. Multiple imputation (MI) with a tipping point approach will be used, i.e. by replacing each missing value with a set of plausible values that represent the uncertainty about the right value to impute, and then applying a series of increasing penalties to the week 4 imputed values in the APL-130277 arm (see Section 15.1.4.1).

For safety data, if severity or causality is missing for an adverse event then it will it be assigned severe or related, respectively.

Mitigation of COVID-19 Impact

Due to the COVID-19 pandemic, sites could deviate from protocol v 4.01 (e.g. missed study visits or scheduled assessments). In accordance with the guidelines released by regulatory health authorities, instructions regarding how a subject's visits and assessments should be handled were sent to the sites. Any notable deviation from protocol v 4.01 due to COVID-19 will be captured and evaluated in the CSR. A memo was issued to study sites for COVID-19 guidance on study procedures. COVID-19 related unscheduled visits and missing data will be handled in multiple ways:

- A listing for all COVID-19 related protocol deviations will be created (section 10).
- Handling of patients requiring dose re-confirmation in Part A is pre-specified (section 7).
- Missing data in efficacy endpoints due to COVID-19 will be handled as described below (sections 15.1.2, 15.1.4, 15.2.2, 15.3.2).
- COVID-19 is one of the reasons for screen failure and study discontinuation and will be summarized in disposition tables.

- COVID-19 related AEs will be captured in EDC and summarized along with other AEs.

8.4. MULTIPLE COMPARISONS/ MULTIPLICITY

The efficacy analyses will be performed using the PART B mITT Population.

The primary and secondary efficacy endpoints will be tested in hierarchical order to maintain an overall type I error rate of 0.05 (i.e. tested in a fixed sequential manner).

The primary endpoint, the mean change from pre-dose to 90 minutes post-dose in MDS-UPDRS Part III Motor Examination score at Week 4 for each treatment in PART B, will be tested first at the 5% significance level. If there is a statistically significant treatment difference, the first secondary endpoint will be tested and declared statistically significant if the p-value is less than 0.05. The testing will continue in this manner for each secondary efficacy endpoint in the hierarchical order shown below:

1. Durability of effect, defined as an Investigator confirmed full "ON" within 30 minutes post-dose and at 90 minutes post-dose, after 4 weeks of dosing in each crossover period (assessed by the blinded-rater in-clinic at V3 and V6 of PART B);
2. Subject preference for APL-130277 treatment as recorded for question 9 of the TPQ (Question 9a or 9b). This assessment is scheduled to be performed after the subject has completed both APL-130277 and s.c. apomorphine treatment regimens (assessed in clinic at V6 of PART B);
3. Subject confirmed durability of effect, defined as subject confirmed full "ON" within 30 minutes post-dose and at 90 minutes post-dose, after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B);
4. Patient Global Impression of Change (PGI-C): Subject improvement of "OFF" episodes, defined as very much better, much better or a little better after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B).

All other efficacy analyses are considered to be exploratory and no other adjustments for multiplicity are planned. Any exploratory p-values will be presented descriptively without formal inference.

8.5. EXAMINATION OF SUBGROUPS

Subgroup analyses will be conducted as exploratory analyses without adjustment for multiplicity. It should be noted that the study was not designed to detect treatment differences with high statistical power within subgroups.

At least the following subgroup analyses have been pre-planned.

The subgroup analyses will be performed for analysis of the primary endpoint. Subgroups with too few subjects per treatment arm (e.g. less than 10) may not be analyzed, for the Part B mITT population.

  - Non-elderly (<65 years) versus elderly (≥65 years) subjects
  - Male versus female subjects
  - PD duration (≤ 10 years versus > 10 years)
  - sc apomorphine naïve vs non-naïve subjects at screening
- Country (Group 1 Germany and Austria; Group 2 UK; Group 3 all the rest of countries including Italy, France, and Spain).
- Total daily levodopa dose at baseline (>=900 mg versus <900 mg).

The following subset will be analyzed separately for demographic and baseline characteristics and key safety outcomes (all SOC/PT AEs table) during the Part B:

- Dose level of APL-130277 (10, 15, 20, 25 or 30 mg) and sc apomorphine (2, 3, 4, 5, 6 mg)
  - For the purpose of this analysis this dose level is defined as the highest dose level of APL-130277/sc apomorphine received during the study Part B.

The following subset will be analyzed separately for demographic and baseline characteristics including disease history during the dose titration phase:

- Subjects randomized to Part A but not Part B

Nausea/vomiting rates will be summarized by anti-emetic use during the titration phase.

9. DISPOSITION AND WITHDRAWALS

The patient disposition will be summarized as follows and presented for each phase, treatment sequence and period, as applicable, and overall.

- Screening
  - The number of subjects screened (i.e. the number of subjects in the ENR)
  - The number (%) of subjects who failed screening (% calculated from the ENR), including the distribution of reasons for failing the screening. In case the subject fails the screening multiple times, all the reasons will be summarized and % will be calculated based on total failure events.
- Titration phase: overall

The percentages will be calculated based on the number of subjects randomized into titration phase, unless otherwise specified.

- The number (%) of subjects who received at least one dose of study treatment during titration (Part A Safety Population); broken down by a cross-tabulation of the highest dose level of APL- 130277 (10 mg, 15 mg, 20 mg, 25 mg or 30 mg and total) and the highest dose level of sc apomorphine (2 mg, 3 mg, 4 mg, 5 mg or 6 mg and total).
- The number (%) of subjects in the Part A mITT Population
- The number (%) of subjects who completed the titration
- The number (%) of subjects who discontinued the study prematurely during Part A and Part B, reason for discontinuation by highest dose level of APL- 130277 received during the titration phase (10 mg, 15 mg, 20 mg, 25 mg or 30 mg and total) and also by highest sc (2 mg, 3 mg, 4 mg, 5 mg, 6 mg, and overall). Subject will be counted under the last treatment/dose level that was received prior to discontinuation.

- Treatment phase: overall and by periods

The percentages will be calculated based on the number of subjects randomized into Part B, unless otherwise specified.

- The number (%) of subjects randomized to treatment phase.
- The number (%) of subjects who received at least one dose of study treatment during the cross-over phase (Part B Safety Population); broken down by a cross-tabulation of the highest dose level of APL- 130277 (10 mg, 15 mg, 20 mg, 25 mg or 30 mg and total) and the highest dose level of sc apomorphine (2 mg, 3 mg, 4 mg, 5 mg or 6 mg and total)
- The number (%) of subjects in Part B mITT Population
- The number (%) of subjects who completed the treatment phase
- The number (%) of subjects who discontinued the study prematurely during the treatment phase, presented by highest dose level of APL- 130277 received during the treatment phase (10 mg, 15 mg, 20 mg, 25 mg or 30 mg and total) and also by highest sc (2 mg, 3 mg, 4 mg, 5 mg, 6 mg, and overall), including the distribution of reasons for discontinuations

Reason for discontinuation will be presented as a stacked histogram, with y-axis of number of subjects who discontinued and x-axis displaying treatment and visit week (0, 2, 4).

10. IMPORTANT PROTOCOL DEVIATIONS

Protocol deviations will be identified and documented during the conduct of the study. Protocol deviations will be categorized by type, and whether major or minor based on clinical review. During the blinded data review meeting (prior to data base lock), the team will review programmed data listings of potential IPDs to identify the “important” protocol deviations (IPDs). Additional IPDs may be identified from review of PD log or other data source. The IPD identification plan will be described in a separate document. All individual IPDs will be presented in a data listing. The number

and percentage of subjects who reported important protocol deviations will be summarized by type of deviation and treatment group for Parts A and B.

All deviations from protocol v 4.01 due to COVID-19 will be presented in a separate listing.

11. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

Demographics and baseline characteristics will be summarized descriptively by treatment sequence, for the Part A Safety population, Part B Safety population, and Part B mITT population.

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 subjects with a score of 30, 29, 28, 27, 26 or <26)
- Modified Hoehn and Yahr scale in "ON" state
- MDS-UPDRS Part I Score at SV2 prior to levodopa dosing
- MDS-UPDRS Part II Score at SV2 prior to levodopa dosing
- MDS-UPDRS Part III Score assessed prior to levodopa dosing at SV2
- MDS-UPDRS Part III Score and change assessed at all timepoints after levodopa dosing at SV2
- Total daily levodopa dose, number of levodopa doses per day.
- MDS-UPDRS III score change from pre-dose to 15, 30, 60, 90 and 120 minutes at the final dose (Part B week 0 dose) during titration- only for Part B mITT population.
- Full "ON"/Partial "ON"/ "OFF" status at 15, 30, 60, 90 and 120 minutes at the final dose (Part B week 0 dose) during titration- only for Part B mITT population, for both investigator and subject confirmed ON/OFF status

  For example, suppose a subject received doses up to 30 mg during titration but received 25 mg in Part B Week 0. Then the final dose for the subject will be 25 mg. The MDS-UPDRS III summary statistics will include this subject's last observed data for 25 mg dose during titration phase.

No statistical testing will be carried out for demographic or other baseline characteristics.

12. MEDICAL HISTORY

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA

v 21.0). The medical history data will be summarized with frequencies and percentages of subjects with at least one medical history item, and subject 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 in Part A, 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) will be summarized.

The summary of medical and PD history data will be done for the Part A Safety population and Part B safety population by treatment sequence and overall.

13. MEDICATIONS

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 Drug Global March 2018 B3).
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 in Part A will be considered concomitant medications. The prior and concomitant medications will be summarized separately.
Summaries of prior and concomitant PD treatment medications (medications with CMCAT = "PARKINSON'S DISEASE MEDICATION") 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 subjects receiving medications will be presented by treatment group and overall for concomitant medications and by treatment sequence and overall for prior medications. 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).

Concomitant medication start and stop date will be taken into consideration to decide the assignment of medication to Part A/Part B and the treatment within the Part. Concomitant medication use will be assigned to each treatment the use coincides with. Hence, a medication maybe assigned to both Parts (Part A and Part B) and/or both treatments in the Part. For example, a concomitant medication that starts in Part A period 1 and ends after start of Part A period 2, will be assigned to both treatment arms in Part A.

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

Also, based on the Part B home diary, the incidence of anti-nausea medication use (for each subject, the percentage of days where nausea medication is recorded out of all days where diary is recorded) will be summarized.

The summary of concomitant medications will be done for the Part A safety population and Part B safety population.

14. STUDY TREATMENT EXPOSURE

Part A exposure for Part A safety population:

The number of subjects by the dose level of APL-130277 (10 mg, 15 mg, 20 mg, 25 mg, or 30 mg) and sc apomorphine (2 mg, 3mg, 4, mg, 5mg, or 6 mg) will be summarized for following:-

- actual dose level
- highest dose received
- final dose for Part B randomization

Part B exposure for Part B safety population:

The following will be summarized by treatment: -

- Subjects exposed to actual treatment
- Duration of exposure (days)

  If period 1 treatment, duration = one day prior to period 2 dose or last dose date (if terminate before period 2) – date of first dose received in period 1 + 1

If period 2 treatment, duration = End of study visit – date of first dose received in period 2 +1

The following will be summarized for home diary:

- Overall diary compliance: number of diaries returned/number of diaries expected (number of days between week 0 and week 4).
- Proportion of subjects using >5, 1-5, or 0 doses/day at least once (during the days when the information was collected). A subject may appear in more than one response but only once in a response for example if out of the 10 diary days, subject took 0 doses on 2 days and 1-5 doses on 8 days. Subject will appear once in response 1-5 doses/day and once in 0 doses/day. Subject is considered non-compliant if they take more than 5 doses in a day.
- Average number of daily doses per day: From all the available diaries for the subject, average number of doses per day for the subject will be calculated (including days with 0 dose taken) and summary statistics presented.
- Average total daily dose (mg) per day: From all available diaries for the subject, total daily dose per day for the subject for each day will be calculated. The total daily dose per day will be averaged across all available days (including days with total daily dose of 0 mg) for the subject and summary statistics presented for the same.
- Percentage of diary responses for the 'Dose Giver', be it the subject or caregiver.

- Final dose for Part B randomization for APL (10 mg, 15 mg, 20 mg, 25 mg, or 30 mg) will be cross tabulated against final dose for Part B randomization for sc apomorphine (2 mg, 3mg, 4, mg, 5mg, or 6 mg)

15. EFFICACY OUTCOMES

15.1. PRIMARY EFFICACY

15.1.1. PRIMARY EFFICACY ENDPOINT & DERIVATION

The primary endpoint is the change from pre-dose to 90 minutes post-dose in MDS-UPDRS Part III Motor Examination score. The primary endpoint is evaluated at V3 and V6, after 4 weeks of dosing of each treatment in each crossover period of PART B, and a blinded rater will be used for this assessment.

The MDS-UPDRS motor score will be calculated as the sum of the individual items of the MDS-UPDRS Part III (items 3.1 – 3.18) and will be obtained separately at each assessment time point.

15.1.2. MISSING DATA METHODS FOR PRIMARY EFFICACY ENDPOINT

Any data collected after rescue medication will not be used for imputation, summary statistics or in statistical model. Subjects who discontinue from the study drug will not be followed and no efficacy data will be collected after discontinuation. MDS-UPDRS III motor score during Part B is to be collected at pre-dose, 15, 30, 60, 90, and 120 minutes post-dose. MDS-UPDRS III is a total of 18 questions with 33 individual items, each item ranges from 0-4. Hence, the MDS-UPDRS III motor score ranges from 0-132. Missing individual items will be imputed using the 2 available values at time points adjacent to the missing item on the same date. For example, for a visit, if an individual item is missing for 30 minutes and 60 minutes, then 30 minutes and 60 minutes can be imputed using available score for 15 minutes and 90 minutes. The maximum of the 2 adjacent values will be assigned as the score for the missing 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. 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 III motor score will be calculated after imputation of the missing item(s) as described above.

If the pre-dose week 0 value is missing (MDS-UPDRS III motor score) it will be imputed using last observed pre-dose value, including screening.

15.1.3. PRIMARY ANALYSIS OF PRIMARY EFFICACY ENDPOINT(S)

15.1.3.1. Estimand

The primary efficacy estimand is defined as the change from pre-dose to 90 minutes post-dose after 4 weeks of dosing, in MDS-UPDRS Part III motor examination score (MDS-UPDRS III motor score)

in levodopa responsive subjects with Parkinson's disease (PD) complicated by motor fluctuations as characterized by the study inclusion/exclusion criteria. The comparison between APL-130277 and sc apomorphine is conducted in the hypothetical setting where subjects were able to stay on study treatment without rescue medication use for 4 weeks.

The five attributes of the primary efficacy estimand are as follows:

A. Population of interest: Levodopa responsive subjects with PD complicated by motor fluctuations, as characterized by the inclusion/exclusion criteria of the study. For the efficacy analyses, the part B mITT population will be used to represent the population of interest.

B. Variable (or endpoint) of interest: Change from pre-dose to 90 minutes post-dose in MDS-UPDRS III motor score after 4 weeks of dosing (Week 4 visit)

C. Treatment: sl APL-130277 (experimental treatment of interest) vs sc apomorphine. The efficacy of sl APL-130277 compared to sc apomorphine as a treatment of "OFF" episodes will be evaluated.

D. Intercurrent event: The intercurrent events that are deemed to have an impact on the interpretation of the endpoint of interest are early withdrawal from study treatment for any reason and rescue medication. These intercurrent events will be handled with the hypothetical strategy. That is, the treatment effect of interest concerns the outcomes had all subjects completed 4 weeks of study treatment without rescue medication use. The efficacy data after the last on-treatment visit or after rescue medication on a visit day are not considered relevant to the treatment effect of interest. For a dose administered at an in-clinic visit in Part B, time of rescue medication taken will be compared with time of MDS-UPDRS III assessment. If the MDS-UPDRS III is assessed after time of rescue medication, that assessment and any assessment after that will be excluded from the analysis. Rather, these data will be implicitly predicted in the mixed-effects model analysis based on assumptions about how the data would have evolved in the absence of treatment withdrawal or administration of rescue medication.

E. Population-level summary for the variable: The estimated LS Mean treatment difference in the change in MDS-UPDRS III motor score from pre-dose to 90 minutes post-dose after 4 weeks of treatment.

15.1.3.2. Justification for the estimand

The primary efficacy estimand defining the treatment effect of interest uses the hypothetical strategy specified in the International Conference on Harmonization (ICH) E9 (R1) Addendum. The primary objective of the study is to demonstrate efficacy of APL-130277 as a treatment of "OFF" episodes compared to sc apomorphine in levodopa responsive subjects with Parkinson's disease (PD) complicated by motor fluctuations. The estimand, or target of estimation, following the hypothetical strategy, is the pharmacological effect seen, had no withdrawals from study treatment or rescue medication use occurred. APL-130277 and sc apomorphine are designed for intermittent PRN use, for the treatment of PD OFF symptoms, and are not intended to be disease modifying treatments. Collection and use of off-treatment/post-rescue data would be difficult to interpret and not clinically relevant in this case. This means that any observations taken after subjects stop study treatment or

receive rescue medication will most likely not contribute relevant information about the pharmacological effect of the drug. Under the hypothetical strategy, the primary endpoint of the trial could be considered as a combination of the observed responses at Week 4 from on-treatment completers and the implicitly model predicted responses at Week 4 for subjects who withdraw from study treatment/receive rescue medication during the trial based on certain assumptions about how the unobserved efficacy outcomes would evolve in the hypothetical setting of no treatment withdrawal/rescue medication administration. However, as such assumptions cannot be directly verified, a tipping point sensitivity analysis will be conducted to assess the robustness of the primary conclusions. No other inclusion/exclusions are defined for this study and any other protocol deviations are ignored in the planned mITT analyses.

15.1.3.3. Hypotheses and statistical model

The primary objective of this study is to show that APL-130277 is superior to sc apomorphine in improving motor function, assessed as the mean change from pre-dose in MDS-UPDRS Part III motor score at 90 minutes after dosing at the 4 week visit (V3 and V6) of the Part B. That is, the null-hypothesis to be tested can be generally stated as follows:

H_0: APL-130277 is the same as sc apomorphine in its effect on the motor function as measured by mean change from pre-dose in MDS-UPDRS III motor score at 90 minutes after 4 weeks in the APL-130277 treatment compared to sc apomorphine.

against the two-sided alternative

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

The primary endpoint will be analyzed in the PART B mITT population and compared between the treatment groups using a linear mixed model (SAS Mixed procedure), as described by Tao et al. (2015). The mixed model for the change from pre-dose to 90 minutes in the MDS-UPDRS III motor score as the response outcomes, includes the following variables: treatment group, visit week (0, 2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. The week 0 visit pre-dose MDS-UPDRS III motor score will be used as a covariate. Subject nested within the PART B sequence will be included as a random effect and an AR(1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The least square (LS) mean, standard error, and LS mean treatment difference, along with the 95% confidence interval (CI) and p-value will be provided. The example SAS code for the analysis is outlined below:

```
proc mixed data=&data;
    class treatment week period sequence subject;
    model UPDRS III pre-dose change = pre-dose treatment week treatment*week sequence period/
ddfm=kr;
    random subject(sequence);
    repeated week / subject=subject *treatment  type=AR(1);
```

```
lsmeans treatment*week /cl alpha=0.05;
run;
```

15.1.4. SENSITIVITY AND SUPPLEMENTARY ANALYSIS OF PRIMARY EFFICACY ENDPOINT

15.1.4.1. Sensitivity Analysis

If the primary analysis significantly favors APL-130277, a missing not at random (MNAR) sensitivity analysis of the primary endpoint will be conducted in the PART B mITT population. This analysis will use a tipping point approach with multiple imputation (MI) for the missing outcome data.

Sensitivity to departures from the MAR assumption will be investigated using this tipping point analysis. In this analysis, departures from MAR in the APL-130277 group only will be assessed assuming that subjects who discontinue the treatment prior to week 4 or who take rescue medication at week 4, prior to 90 minutes post-dose, have efficacy outcomes after discontinuation/rescue medication at week 4 that on average are worse by some amount δ compared to other similar subjects with observed data at the same time point or compared to a value which would have been assumed under a MAR model.

A series of analyses will be performed with increasing values of δ until the analysis conclusion of a statistically significant treatment effect no longer holds. Successively increasing deltas will be imposed on the imputed values at Week 4 in the APL-130277 group only, starting with an increment (worsening) of 1.0 point. The delta will be further increased in steps of 0.5 points (1.5, 2.0, 2.5…) until the statistical significance is lost. For the sc apomorphine group, the MI using the MAR assumption will be used for missing data. The value of δ that overturns the primary results will represent a tipping point, i.e. the point at which the p-value becomes >0.05. An evaluation of the clinical plausibility of the tipping point value will be provided.

Intermittent (non-monotone) missing data (MDS-UPDRS III motor score change from pre-dose to 90 minutes) will be imputed based on the MAR assumption using the Markov Chain Monte Carlo (MCMC) method to obtain datasets with a monotone missing pattern.

The remaining monotone missing data will then be imputed using sequential regression multiple imputation, where a separate regression model is estimated for imputation of each variable (i.e., change from pre-dose to 90 minutes in MDS-UPDRS III motor score at each visit week). Each regression model will include explanatory variables for all previous visit weeks (0, 2, 4) change from pre-dose to 90 minutes in MDS-UPDRS III motor score and treatment. Treatment periods will be considered independent for the purposes of imputation.

After the MAR-based imputations have been generated for change from pre-dose to 90 minutes in MDS-UPDRS III motor score at each visit, the imputed values for the change from pre-dose to 90 minutes MDS-UPDRS motor score at week 4 in the APL-130277 group will be penalized by a value

of delta.

Five hundred (500) imputed datasets will be generated. The random seed numbers 12345 is for the partial MCMC imputation step and 56789 is for the sequential regression, monotone imputation step.

Each of the 500 imputed and delta-adjusted datasets will be analyzed using an ANCOVA model for week 4. The ANCOVA model would include the following variables: treatment group, PART B sequence, and period as fixed effects. The week 0 pre-dose UPDRS III score will be used as a covariate. Subject nested within sequence will be included as a random effect. Results from the analysis of each imputed dataset, i.e. the LS mean treatment difference and its standard error, will be combined using Rubin's imputation rules (using the SAS MIANALYZE procedure) to produce a pooled LS mean estimate of treatment difference, its standard error and 95% CI, and a pooled p-value for the test of null hypothesis of no treatment effect.

Analyses will be conducted with different values of delta, starting with a change in pre-dose to 90 minutes in MDS-UPDRS Part III motor score with an increment (worsening) of 1.0 point. The delta is further increased in the steps of 0.5 points (1.5, 2.0, 2.5 …) until the statistical significance is lost, i.e. until the p-value becomes >0.05.

15.1.4.2. Supplementary Analyses

Two supplementary analyses of the primary endpoint will be conducted.

- The first supplementary analysis is an analysis of just week 4 outcome data using ANCOVA model. The ANCOVA model would include the following variables: treatment group, PART B sequence, and period as fixed effects. The week 0 visit pre-dose MDS-UPDRS III motor score will be used as a covariate. Subject nested within sequence will be included as a random effect. Missing Week 4 data within each period will be imputed using LOCF for this analysis (i.e. carrying-forward available week 0 or 2 change from pre-dose data). If there are convergence or other numerical issues that prevent fitting the primary repeated measures model analysis specified above, this supplementary ANCOVA analysis will be used instead for primary inference.
- The second supplementary analysis is a binary responder analysis. A responder is defined as a subject with a > 4 point improvement in the change from pre-dose to 90 minutes post-dose in MDS-UPDRS Part III Motor score. Response will be compared between the treatment groups in the PART B mITT population using a generalized linear random effects model with logit link function. The model includes the binomial variables as the response outcomes, and the following independent variables: treatment group, visit week (0, 2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. Subject nested within the PART B sequence will be included as a random effect and an AR(1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The SAS Glimmix procedure will be used for

this analysis. The predicted response rates, treatment odds ratio, 95% confidence interval and p- value will be presented. The example SAS code for the analysis is outlined below:

```
Proc glimmix data=&data;
class treatment week sequence period subject;
model response=treatment week treatment*week sequence period/dist=binary link=logit ddfm=kr;
random subject(sequence);
random week/subject=treatment*subject type=AR(1) residual;
lsmeans treatment*week/cl ilink diffs oddsratio;
run;
```

15.1.4.3. Supportive Analysis of Primary Efficacy Endpoint

Observed and change from pre-dose MDS-UPDRS III motor scores will be summarized descriptively by visit and time point for the PART B mITT population.

A graph will be generated using the estimates calculated from mixed models: the estimated LS mean changes (with SEM) from pre-dose to 90 minutes post-dose by treatment group and visit week. The x-axis includes the visit week (week 0, week 2, or week 4) and the y-axis the change from pre-dose.

15.2. SECONDARY EFFICACY

The secondary efficacy analyses will be performed for the Part B mITT Population, unless specified otherwise.

15.2.1. SECONDARY EFFICACY ENDPOINTS & DERIVATIONS

15.2.1.1. Investigator confirmed durability of Effect

The first secondary efficacy is defined as the investigator confirmed full "ON" within 30 minutes post-dose (i.e. "ON" at either 15 minutes and/or at 30 minutes) and at 90 minutes post-dose, after 4 weeks of dosing in each Part B crossover period.

A subject would have a value for endpoint if they do have a valid assessment within 30 minutes and at 90 minutes post dose. For example, suppose a subject turns ON at either 15 minutes or 30 minutes and is ON at 90 minutes on V1, V2. Also, subject is ON at 30 and OFF at 90 min at V3 and V4. For example, suppose subject receives rescue mediation at 45 minutes on V5, and discontinues prior to V6. The responder endpoint values for this subject for all visits will be: yes, yes, no, no, blank, blank. No imputation will be done for missing values.

15.2.1.2. Subject Treatment Preference

The second secondary preference endpoint is subject treatment preference based on question 9a (Q9a) or 9b (Q9b) of the TPQ using the PART B mITT population. Q9a is assessed on a 5-point Likert scale and Q9b is assessed as a VAS score. Q9a responses will be dichotomized as follows for statistical analysis: preference for APL (responses of either definitely or somewhat prefer APL) versus no preference for APL (responses of no preference, or somewhat/definitely prefer sc apomorphine). Prior to amendment 2, subject preference was assessed using a visual analogue scale (VAS) score (-50 to 50). The VAS score will also be dichotomized as preference for APL-130277 (score of >0 to 50) versus no preference for APL-130277 (-50 to 0). The VAS dichotomized values based on TPQ: Q9b and subject preference based on TPQ: Q9a dichotomized values will be combined. If a subject responded to both Q9a and Q9b, then Q9a only will be used for combined analysis.

The preference assessment is scheduled to be performed upon completion of both treatments in PART B. For subjects who terminate the study early, the assessment at the early termination (ET) visit will be used for analysis if available. If in this case, the ET assessment is not available, the last treatment received prior to termination will be considered the treatment that is not preferred (i.e. the alternative treatment, whether received or not, will be considered preferred for analysis).

15.2.1.3. Subject confirmed durability of effect

The third secondary efficacy is defined as the subject confirmed full "ON" within 30 minutes post-dose (i.e. "ON" at either 15 minutes and/or at 30 minutes) and at 90 minutes post-dose, after 4 weeks of dosing in each Part B crossover period. No imputation will be done for missing values.

15.2.1.4. Patient Global Impression of Change (PGI-C)

The fourth secondary endpoint based on PGI-C scale is defined as the subject improvement of "OFF" episodes, defined as very much better, much better or a little better at Week 4 in each PART B crossover period. If the Week 4 PGI-C assessment is missing, the assessment at the ET visit for that treatment period will be used if available for analysis.

15.2.2. MISSING DATA METHODS FOR SECONDARY EFFICACY ENDPOINT(S)

No imputation will be done for durability of effect.

The preference assessment is scheduled to be performed upon completion of both treatments in PART B. For subjects who terminate the study early, the assessment at the ET visit will be used for analysis if available. If in this case, the ET assessment is not available, the last treatment received prior to termination will be considered the treatment that is not preferred (i.e. the alternative treatment,

whether received or not, will be considered preferred for analysis).

If the Week 4 PGI-C assessment is missing, the assessment at the ET visit for that treatment period will be used, if available for analysis.

15.2.3. PRIMARY ANALYSIS OF SECONDARY EFFICACY ENDPOINTS

15.2.3.1. Analysis of Investigator and subject confirmed durability of effect

For these endpoints, response will be compared between the treatment groups in the PART B mITT population using a generalized linear random effects model with logit link function. The model includes the binomial variables as the response outcomes, and the following independent variables: treatment group, visit week (0, 2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. Subject nested within the PART B sequence will be included as a random effect and an AR (1) covariance structure will be used for the repeated measures over time (visit week). The KR method will be used to calculate degrees of freedom. The SAS Glimmix procedure will be used for this analysis. The predicted response rates, treatment odds ratio, 95% confidence interval and p-value will be presented. The SAS code planned for the analysis is similar to the responder analysis described in section 15.1.4.2. This method of analysis will be utilized for the following secondary endpoints:

- Investigator confirmed durability of Effect
- Subject confirmed durability of effect

The following graphs will be generated using the estimates calculated from the model

- Predicted response rate for investigator confirmed durability of effect by treatment group and visit week. The x-axis includes the visit week (week 0, week 2, or week 4) and the y-axis the predicted response rate
- Predicted response rate for subject confirmed durability of effect by treatment group and visit week. The x-axis includes the visit week (week 0, week 2, or week 4) and the y-axis the predicted response rate

15.2.3.2. Subject Treatment Preference

Subject reported preference for APL-130277 or sc apomorphine is based on question Q9a or Q9b of the TPQ. The number and percent of subjects in each response category will be summarized descriptively. The main analysis will evaluate the proportion of subjects with preference for APL.

The proportion of subjects preferring the APL-130277 treatment and two-sided 95% confidence interval for this proportion will be calculated. The p-value from a 1-sample, 2-sided test of the null

hypothesis that the true proportion is 50% will be calculated to evaluate if a significantly higher proportion of the subjects prefer APL-130277 or not. The confidence interval and p-value will be calculated using the binomial distribution with normal approximation (Wald asymptotic confidence interval).

15.2.3.3. Patient Global Impression of Change

PGI-C will be analyzed using a generalized linear random effects model with logit link function. The model includes the following independent variables: treatment group, PART B sequence, and period as fixed effects and PGI-S at V1 as binary covariate (with normal/mild (0) vs moderate/severe (1)). Subject nested within sequence will be included as a random effect. The predicted response rates, treatment odds ratio, 95% confidence interval and p-value will be presented.

15.2.4. SUPPLEMENTARY ANALYSIS OF SECONDARY EFFICACY ENDPOINTS

15.2.4.1. Supplementary Analysis

As a supplementary analysis, preference will be summarized on the basis of TPQ: Q9a and Q9b separately. Responses will be dichotomized as described above in section 15.2.1.2.

As a supportive analysis for the subject preference endpoint (Q9a and Q9b combined), the PART B sequence effect (the order in which the treatments were given in PART B) will be assessed by summarizing the proportion of subjects preferring the APL-130277 treatment within each of the two sequences.

15.3. OTHER EFFICACY ENDPOINTS (PART B)

15.3.1. OTHER EFFICACY ENDPOINTS & DERIVATIONS

15.3.1.1. Clinical Global Impression of Improvement (CGI-I)

Subject improvement, defined as very much improved, much improved, or minimally improved after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B)

The proportions of subjects who responded to CGI-I as very much improved, much improved, or minimally improved will be defined as improved.

For subjects who terminate the study early, the assessment at the early termination (ET) visit will be used for analysis if available. The ET assessment will be assigned to the week 4 for the last treatment received prior to discontinuation.

15.3.1.2. Change from pre-dose to 15, 30, 60, 90, and 120 minutes post-dose in MDS-UPDRS Part III score (assessed in-clinic V1, V2, V4, and V5). Change from pre-dose to 15, 30, 60 and 120 minutes post-dose in MDS-UPDRS Part III score (assessed in-clinic V3 and V6)

MDS-UPDRS Part III scores will be derived as described in primary efficacy section 15.1.

15.3.1.3. Investigator confirmed full "ON" at the 15, 30, 60, 90, and 120 minutes post-dose time points (assessed in-clinic V1, V2, V3, V4, V5, and V6)

At each week (week 0, 2, and 4) and at timepoints 15, 30, 60, 90 and 120 minutes post-dose, the investigator will confirm the "investigator rating of OFF and ON" i.e., if the subject was "ON", "OFF" or "Partially ON" at each timepoint. For this analysis, the subject will be considered full "ON" if the confirmed investigator rated assessment of "ON" is recorded on the CRF.

15.3.1.4. Subject confirmed full "ON" at the 15, 30, 60, 90, and 120 minutes post-dose time points (assessed in-clinic V1, V2, V3, V4, V5, and V6)

At each week (week 0, 2, and 4) and at timepoints 15, 30, 60, 90 and 120 minutes post-dose the investigator will confirm the "subject rating of OFF and ON', i.e. if the subject was "ON", "OFF" or "Partially ON" at each timepoint. For this analysis, the subject will be considered full "ON" if the confirmed subject rated assessment of "ON" is recorded on the CRF.

15.3.1.5. Time to full "ON" and Time to partial "ON" as determined by the subject and Investigator (assessed in-clinic at all visits in PART B)

Time to full "ON" will be calculated by evaluating the difference in time from the "please indicate the time of first full ON after dosing (collected on "investigator rating of OFF and ON" CRF) and the "Time the sublingual thin film was placed under the tongue" (collected on "in-clinic dosing" CRF).

If "please indicate the time of first full ON after dosing" is left blank or marked as "N/A", then time to full "ON" will be calculated per the confirmation of "FULL ON" at 15, 30, 60, 90 and 120 minutes after dosing (collected on "investigator rating of OFF and ON" CRF).

If the subject did not turn "ON" after dosing, the subject will be considered censored at 120 minutes after dosing (the last assessment time)

For a blank value or "N/A" ticked for the question "Time the sublingual thin film was placed under the tongue" and/or Time to full "ON" not available from "investigator rating of OFF and ON", the data will be treated as having a missing value.

This will be calculated for both the investigator rating and subject rating (based on "subject rating of

OFF and ON" CRF).

Time to partial "ON" will be calculated similarly.

15.3.1.6. Per the Expanded Home Dosing Diary, percent of episodes with a subject-rated full "ON" within 30 and at 90 minutes post-dose based on the 3 consecutive days prior to V2, V3, V5, and V6

During the Part B of the study, subjects will complete the expanded home dosing diary on the 3 consecutive days prior to their next scheduled in-clinic visit. The subjects will self-administer their doses in order to treat up to 5 "OFF" episodes per day. The subjects will fill in the time when study treatment is self-administered, "ON" within 30 minutes of dosing status and "ON" at 90 minutes after dosing. In total, each subject can record up to 15 episodes at 2 visits after the randomization for Part B (visit week 2 and 4), i.e. a total of up to 30 episodes. For each subject, the percentage of episodes in which the full "ON" response was achieved within 30 minutes and was "ON" at 90 minutes, out of all recorded episodes will be calculated. In case the time of self-administration has been recorded but the corresponding "ON" status for either within 30 minutes and/or 90 minutes is missing, the status will be classified as non-response for the numerator but will be counted in the total recorded episodes. The percentages calculated separately for each subject prior to each visit (week 2 and 4) will be used as response variables in the statistical analysis. For subjects who did not record any episodes during the part B phase, the endpoint will be set as missing.

15.3.1.7. Level of satisfaction as assessed by the Treatment Satisfaction Questionnaire for Medication (assessed in-clinic at V3 and V6)

The treatment satisfaction questionnaire is a set of 14 items with multiple choices. The assessment will evaluate the effectiveness, side effects, and convenience of the medication over the previous 4 weeks.

For subjects who terminate the study early, the assessment at the early termination (ET) visit will be used for analysis if available. The assessment at ET will be assigned to the week 4 for the last treatment received prior to discontinuation.

15.3.1.8. MDS-UPDRS – Part I Score: change from Screening to V2, V3, V5, and V6 (assessed in-clinic at V2, V3, V5, and V6)

Part I (Non-Motor Aspects of Experiences of Daily Living) of the MDS-UPDRS will be completed at Screening and at PART B V2, V3, V5, and V6. The MDS-UPDRS Part I score will be calculated as the sum of the individual items of the MDS-UPDRS Part I questionnaire (items 1.1 – 1.13). Missing individual items will not be imputed. If there is at least 1 missing item, the corresponding MDS-UPDRS Part I score will be set as missing.

15.3.1.9. MDS-UPDRS – Part II Score: change from Screening to V2, V3, V5, and V6 (assessed in-clinic at V2, V3, V5, and V6)

Part II (Motor Aspects of Experiences of Daily Living) of the MDS-UPDRS will be completed at Screening and at PART B V2, V3, V5, and V6. 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.

15.3.1.10. MDS-UPDRS – Part IV Score: change from Screening to V2, V3, V5, and V6 (assessed in-clinic at V2, V3, V5, and V6)

Part IV (Motor Complications) of the MDS-UPDRS will be completed at Screening and at PART B V2, V3, V5, and V6. 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.

15.3.1.11. Improvement in amount of troublesome dyskinesia in the last month per the in-clinic dyskinesia questionnaire; defined as less than usual amount of troublesome dyskinesia after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6)

At V3 and V6 only, the subject will be asked “Since you started the most recent treatment for your “OFF” episodes four weeks ago, how would you rate the amount of troublesome dyskinesia you had in the last month”. The subject can choose from four options: no troublesome dyskinesia, less than usual, the same as usual, and more than usual. If the subject discontinues the study early, the subject will be assessed at the Early Termination Visit and this will be used for analysis. The ET assessment will be assigned to the last treatment used. If the subjects choose the options of ‘no troublesome dyskinesia’ or ‘less than usual’ then it will be considered as improvement in amount of troublesome dyskinesia. If the status is not chosen, the subject will not be considered to have improved and will be assigned the status of “the same as usual”.

15.3.1.12. Percent of “ON” episodes without troublesome dyskinesia (per expanded home diary) based on the 3 consecutive days prior to V2, V3, V5, and V6

During the Part B of the study, subjects will complete the expanded home dosing diary on the 3 consecutive days prior to their next scheduled in-clinic visit. The subjects will self-administer their doses in order to treat up to 5 “OFF” episodes per day. The subjects will fill in the time when study treatment is self-administered, “ON” within 30 minutes of dosing status and “ON” at 90 minutes after dosing. Also, if the subject reported dyskinesia anytime they were “ON” they will also state if it was troublesome or not. In total, each subject can record up to 15 episodes at 2 visits after the

randomization for Part B (visit week 2 and 4), i.e. a total of up to 30 episodes.

For each subject, the percentage of episodes in which dyskinesia was recorded and was not troublesome, out of all recorded episodes will be calculated. In case the time of self-administration has been recorded but the corresponding answer to the question: did you have dyskinesia anytime whilst you were "ON" and/or was it troublesome? is missing, the status will be classified as non-response for the numerator but will be counted in the total recorded episodes.

15.3.1.13. Change in the PDQ-39 Total Index Score, and Mobility and Activities of Daily Living subscale scores from pre-dose V1 to after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6)

Change from pre-dose V1 to week 4 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 subject 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.
For subjects who terminate the study early, the assessment at the early termination (ET) visit will be used for analysis if available. The assessment at ET will be assigned to the week 4 for the last treatment

received prior to discontinuation.

15.3.1.14. European Quality of Life – 5 dimensions (EQ-5D-5L) (assessed on-clinic at V1, V3 and V6)

The EQ-5D is a utility scale consisting of three components: health state dimensions, health state thermometer scale and health state index.
For subjects who terminate the study early, the assessment at the early termination (ET) visit will be used for analysis if available. The assessment at ET will be assigned to the week 4 for the last treatment received prior to discontinuation.

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 V1 to week 4 on the VAS scores will be evaluated.

15.3.1.15. Ease of Use (assessed on-clinic at V3 and V6)

The ease of use questionnaire has three questions (with 5 response choices)

1. Opening the package (very easy to open, easy to open, neither easy not difficult to open, difficult to open, very difficult to open)

2. Handling (very easy to handle, easy to handle, neither easy not difficult to handle, difficult to handle, very difficult to handle)

3. Dosing (very easy to dose myself, easy to dose myself, neither easy nor difficult to dose myself, difficult to dose myself, very difficult to dose myself)

   For subjects who terminate the study early, the assessment at the early termination (ET) visit will be used for analysis if available. The assessment at ET will be assigned to the week 4 for the last treatment received prior to discontinuation.

15.3.1.16. Analysis for Part A assessments

Change from pre-dose to all post-dose timepoints in MDS-UPDRS III at the final dose during titration (both APL and sc apomorphine) will be summarized for the Part B mITT population.

MDS-UPDRS III motor score, dyskinesia, ON/OFF assessment by investigator and subject, home dosing diary assessment for APL and in-clinic assessment for sc apomorphine during part A. All these assessments will be provided as a listing.

15.3.2. MISSING DATA METHODS FOR EXPLORATORY EFFICACY ENDPOINT(S)

Not applicable.

15.3.3. ANALYSIS OF EXPLORATORY EFFICACY ENDPOINTS

15.3.3.1. Clinical Global Impression of Improvement (CGI-I): Subject improvement, defined as very much improved, much improved, or minimally improved after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6 of PART B)

The proportions of subjects who improved (defined as very much improved, much improved, or minimally improved) will be tabulated by treatment group. 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 based on the observed results. An analysis will be conducted similar to the analysis for PGI-C specified in section 15.2.3.3. CGI-I will be analyzed using a generalized linear random effects model with logit link function. The model includes the following independent variables: treatment group, PART B sequence, and period as fixed effects and CGI-S at V1 as covariate. Subject nested within sequence will be included as a random effect. The predicted response rates, treatment odds ratio, 95% confidence interval and p-value will be presented.

15.3.3.2. Change from pre-dose to 15, 30, 60, 90, and 120 minutes post-dose in MDS-UPDRS Part III score (assessed in-clinic V1, V2, V4, and V5). Change from pre-dose to 15, 30, 60 and 120 minutes post-dose in MDS-UPDRS Part III score (assessed in-clinic V3 and V6)

Missing data for all the timepoints will be imputed as described in section 15.1.2. Similar to the primary endpoint, any data collected after rescue medication will not be used and no efficacy data will be collected after subject discontinues.

Observed and change from pre-dose MDS-UPDRS III motor scores will be summarized descriptively by visit and time point for the PART B mITT population.

The MDS-UPDRS III motor score at the 90 minutes timepoint at each visit week (week 0, 2, and 4) will be evaluated using the primary analysis model. A similar model separately for each of the other four timepoints (15, 30, 60, and 120 minutes) will be fitted. The least square (LS) mean, standard error, and LS mean treatment difference, along with the 95% confidence interval (CI) and p-value will be presented.

A graph of the estimated LS mean changes (with SEM) from pre-dose to 15, 30, 60, 90 and 120 minutes by treatment group will be presented. Separate graphs will be produced for each visit week (0, 2, and 4). The x-axis includes the time of the assessment (15, 30, 60, 90 or 120 minutes) and the y-axis the

LS mean change from pre-dose. LS mean estimates and SEMs of these estimates from the mixed models for each timepoint will be used for the display.

15.3.3.3. Investigator confirmed full "ON" at the 15, 30, 60, 90, and 120 minutes post-dose time points (assessed in-clinic V1, V2, V3, V4, V5, and V6)

The percentage of subjects with an Investigator confirmed full "ON" at each visit week (0, 2, and 4) and time point (15, 30, 60, 90 and 120 minutes) will be summarized descriptively by treatment group. Response will be defined as a subject having investigator confirmed Full "ON". Any data collected after rescue medication will not be used and no efficacy data will be collected after subject discontinues. A separate analysis model for each timepoint similar to the analysis model described in the secondary endpoint in section 15.2.3.1 will be fitted. The model includes the binomial variables as the response outcomes, and the following independent variables: treatment group, visit week (0, 2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. Subject nested within the PART B sequence will be included as a random effect and an AR (1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The SAS Glimmix procedure will be used for this analysis. The predicted response rates, treatment odds ratio, 95% confidence interval and p-value will be presented.

15.3.3.4. Subject confirmed full "ON" at the 15, 30, 60, 90, and 120 minutes post-dose time points (assessed in-clinic V1, V2, V3, V4, V5, and V6)

The percentage of subjects with a subject- rated full "ON" response at each visit week (0, 2, and 4) and time point (15, 30, 60, 90 and 120 minutes) will also be summarized descriptively by treatment group. Response will be defined as a subject-confirmed full "ON". Any data collected after rescue medication will not be used and no efficacy data will be collected after subject discontinues A separate analysis model for each timepoint, similar to the analysis model described in the secondary endpoint in section 15.2.3.1 will be fitted. The model includes the binomial variables as the response outcomes, and the following independent variables: treatment group, visit week (0, 2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. Subject nested within the PART B sequence will be included as a random effect and an AR (1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The SAS Glimmix procedure will be used for this analysis. The predicted response rates, treatment odds ratio, 95% confidence interval and p-value will be presented.

15.3.3.5. Time to full "ON" and Time to partial "ON" as determined by the subject and Investigator (assessed in-clinic at all visits in PART B)

The time to "ON"/partial "ON" at each visit week will be described using the Kaplan-Meier method, including the estimate of the median time to ON and corresponding 95% confidence intervals. Furthermore, the Kaplan-Meier probability estimates at 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115 and 120 minutes will be summarized.

Any data collected after rescue medication will not be used and no efficacy data will be collected after subject discontinues

15.3.3.6. Per the Expanded Home Dosing Diary, percent of episodes with a subject-rated full "ON" within 30 and at 90 minutes post-dose based on the 3 consecutive days prior to V2, V3, V5, and V6

The analysis will be conducted similar to the one used for primary analysis for the primary endpoint. The mixed model for the percent of episodes as the response outcomes, includes the following variables: treatment group, visit week (2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. Subject nested within the PART B sequence will be included as a random effect and an AR(1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The least square (LS) mean, standard error, and LS mean treatment difference, along with the 95% confidence interval (CI) and p-value will be provided.

15.3.3.7. Level of satisfaction as assessed by the Treatment Satisfaction Questionnaire for Medication (assessed in-clinic at V3 and V6)

The treatment satisfaction questionnaire items will be evaluated descriptively by presenting the distribution of responses separately for each of the 14 items. If the subject discontinues the study early, the subject will be assessed at the Early Termination (ET) Visit and this will be used for analysis. The ET assessment will be assigned to the last treatment used.

15.3.3.8. MDS-UPDRS – Part I Score: change from Screening to V2, V3, V5, and V6 (assessed in-clinic at V2, V3, V5, and V6)

The endpoint will be analyzed using a model similar to that of primary analysis model for the primary endpoint section 15.1.3.3. The mixed model for the percent of episodes as the response outcomes, includes the following variables: treatment group, visit week (2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. The screening visit MDS-UPDRS-Part I score will be used as a covariate. Subject nested within the PART B sequence will be included as a random effect and an AR(1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The least square (LS) mean, standard error, and LS mean treatment difference, along with the 95% confidence interval (CI) and p-value will be provided.

15.3.3.9. MDS-UPDRS – Part II Score: change from Screening to V2, V3, V5, and V6 (assessed in-clinic at V2, V3, V5, and V6)

The endpoint will be analyzed using a model similar to that of primary analysis model for the primary

endpoint section 15.1.3.3. The mixed model for the percent of episodes as the response outcomes, includes the following variables: treatment group, visit week (2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. The screening visit MDS-UPDRS-Part II score will be used as a covariate. Subject nested within the PART B sequence will be included as a random effect and an AR(1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The least square (LS) mean, standard error, and LS mean treatment difference, along with the 95% confidence interval (CI) and p-value will be provided.

15.3.3.10. MDS-UPDRS – Part IV Score: change from Screening to V2, V3, V5, and V6 (assessed in-clinic at V2, V3, V5, and V6)

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 screening

15.3.3.11. Improvement in amount of troublesome dyskinesia in the last month per the in-clinic dyskinesia questionnaire: defined as less than usual amount of troublesome dyskinesia after 4 weeks of dosing in each crossover period (assessed in-clinic at V3 and V6)

The distribution of categories (less than usual, the same as usual, and more than usual) will be tabulated by treatment.

15.3.3.12. Percent of “ON” episodes without troublesome dyskinesia (per expanded home diary) based on the 3 consecutive days prior to V2, V3, V5, and V6

The percentages calculated separately for each subject will be summarized descriptively by treatment and week.

The mixed model for the percent of episodes as the response outcomes, includes the following variables: treatment group, visit week (2, 4), treatment by visit week interaction, PART B sequence and period as fixed factors. Subject nested within the PART B sequence will be included as a random effect and an AR(1) covariance structure will be used for the repeated measures over time (visit week). The Kenward-Roger (KR) method will be used to calculate degrees of freedom. The least square (LS) mean, standard error, and LS mean treatment difference, along with the 95% confidence interval (CI) and p-value will be provided.

15.3.3.13. Change in the PDQ-39 Total Index Score, and Mobility and Activities of Daily Living subscale scores from pre-dose V1 to after 4 weeks of dosing in each crossover period (assessed in-clinic at V1, V3 and V6)

Absolute value and Change in PDQ-39 total index score and sub-scores will be summarized descriptively.

15.3.3.14. European Quality of Life – 5 dimensions (EQ-5D-5L) (assessed in-clinic at V1, V3 and V6)

The health state dimensions will be evaluated by presenting the distribution of responses separately for each of the 5 dimensions by treatment and week.

15.3.3.15. Ease of Use Questionnaire (assessed in-clinic at V3 and V6)

The ease of questions, three questions will be evaluated by presenting the distribution of responses separately for each of the questions by treatment.

15.4. SUB-GROUP ANALYSIS

15.4.1. SUB-GROUP ANALYSIS FOR PRIMARY ENDPOINT

Subgroup analyses for the primary endpoints will be performed for factors defined in Section 8.5 of this document. For each of the subgroup factors, a mixed 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 estimated treatment effects from this model will be summarized by subgroup and visit week. The influence of each subgroup factor will also be evaluated using the p-value for the treatment by subgroup interaction term calculated from with this model.

16. SAFETY OUTCOMES

All tables for safety outcomes will be based on the PART A safety Population or Part B safety population, as appropriate.

There will be no statistical comparisons between the treatment groups for safety data.

16.1. ADVERSE EVENTS

Adverse Events (AEs) will be coded using Medical Dictionary for Regulatory Activities (MedDRA), Version 21.0.

Pre-treatment events are untoward medical occurrences that start after the subject has signed the

Informed Consent Form but prior to the first dose of study treatment. Adverse Events (AEs) are untoward medical occurrences that start after the subject receives the first dose of study treatment. PART A AEs are defined as all AEs that start on or after the date of the first dose of study drug (APL-130277 or sc apomorphine) in PART A, but before the date of the first dose of study drug in PART B. PART B AEs are defined as all AEs that start on or after the date of the first dose of study drug in PART B.

Within each part of the study, the AE will be assigned to the study treatment (APL-130277 or sc apomorphine) that was last received on or prior to the day of the AE onset (i.e. the last study treatment that was received before the onset of the AE). AEs with onset after the last dose of the study treatment are attributed to the treatment received during the last treatment period within each study part. Both event and subject counts, where applicable, will be summarized. The counts will be complemented by percentages calculated for the subject counts unless otherwise specified.

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

- 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 include AEs and pre-treatment events.

An overall summary of number of subjects within each of the categories described in the sub-section below, will be provided as specified in the table shell templates.

For all summaries, each subject will be counted only once within each category (e.g. an AE type, a severity level, a relationship level, a SOC, and a PT). For summaries by SOC and PT, AEs will be sorted by SOC alphabetically and by PT in decreasing frequency in the “APL-130277” column. For summaries by SOC and PT, AEs will be sorted by SOC alphabetically and by PT in decreasing frequency in the “APL-130277” column. Only number and percentage of subjects reporting SAEs/AEs will be presented for “Any SAEs/AEs” row and no subject or event count will be presented for SOC.

16.1.1. ALL AEs

The AE tables will be presented for

- PART A tables: present AEs that occur in Part A only and in both Parts A and B combined by study treatment
- PART B tables: present AEs that occur in part B by study treatment

Incidence of AEs will be presented by System Organ Class (SOC) and Preferred Term (PT) and study treatment, and also broken down further by maximum severity and relationship to study treatment.

- An overall summary of the number and percentage of subjects reporting AEs and the number of AEs for all AEs, drug-related AEs, severe AEs, serious AEs, non-serious AEs, AEs leading to study drug withdrawal, AEs leading to dose reduction, AEs leading to drug interruption, and AEs leading to death;
- AEs by SOC and PT, number and percentage of subjects reporting AEs and the number of AEs;
- AEs by SOC and PT, by actual dose received just prior to occurrence of AE number and percentage of subjects reporting AEs and the number of AEs;
- AEs by PT, number and percentage of subjects reporting AEs and the number of AEs;
- AEs by SOC and PT, present all SOC and PT that are at least 5% for APL treatment, by subject count (%) for PT only

16.1.1.1. Severity

Severity is classified as mild/ moderate/ severe (increasing severity). AEs starting after the first dose of study treatment with a missing severity will be classified as severe

If a subject reported an AE more than once within the same SOC/ PT with different severity levels, the subject will be assigned to a severity level for that SOC/ PT based on the worst case severity (i.e. maximum severity). Event counts will not be included in this summary.

This will be summarized, separately for Part A table and Part B table.

16.1.1.2. Relationship to Study treatment

Relationship, as indicated by the Investigator, is classified as “not related”, “possible”, “probable”, or “definite” (increasing strength of relationship). A “related” AE is defined as an AE with a relationship to study treatment as “possible”, “probable”, or “definite” related to study treatment. AEs with a missing relationship to study treatment will be regarded as *“related”* to study treatment.

If a subject reported an AE more than once within the same SOC/ PT in different relationship categories, the subject will be assigned to a category for that SOC/ PT based on the worst case relationship (i.e. strongest relationship). Event counts will not be included in this summary.

16.1.2. AEs Leading to Withdrawal of Study Treatment

AEs leading to permanent discontinuation of study treatment will be identified by using the AE action taken as "Drug Withdrawn" from the AE page of the (e)CRF.

For AEs leading to discontinuation of study treatment, summaries for number and percentage of subjects reporting AEs and the number of AEs by SOC and PT will be prepared, for Part A table and Part B table. Similar tables by actual dose level will also be summarized by Part B. The actual dose level prior to the AE onset from diary outpatient accountability form will be used for this analysis.

16.1.3. SERIOUS AND NON-SERIOUS ADVERSE EVENTS

Serious adverse events (SAEs) are those events recorded as "Serious" on the Adverse Events page of the (e)CRF. A summary of serious AEs by SOC and PT, number and percentage of subjects reporting SAEs and the number of SAEs will be prepared. A similar table by actual dose level, i.e. last prescribed dose prior to SAE, will also be summarized for Part B.

All other events not recorded as "Serious" will be considered as "Non-Serious". All non-serious AEs will be summarized by SOC and PT, and by PTs where at least one treatment has more than 5% of subjects for any PT. Number and percentage of subjects reporting AEs and the number of AEs will be presented for each PT.

16.1.4. ADVERSE EVENTS LEADING TO DEATH

AEs leading to Death are those events which are recorded as "Fatal" on the Adverse Events page of the (e)CRF. A summary of AEs leading to death by SOC and PT, number and percentage of subjects reporting AEs and the number of AEs will be prepared separately for Part A and Part B.

16.1.5. ADVERSE EVENTS OF SPECIAL INTEREST

Separate summaries will be generated for AEs of special interest (AESI), separately for Part A and Part B, tabulated by category and PT using the categories as specified below:
Hypotension, orthostatic hypotension: defined as all AEs with HLGT "Decreased and nonspecific blood pressure disorders and shock"

- Syncope: defined as all AEs 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" (Midrace 10045555)

Or HLGT of "Seizures (incl subtypes)"

- Falls & injuries: defined as all AEs meeting the criteria for the standardized MedDRA Query (SMQ) "Accidents and injuries" (narrow terms)
- Dyskinesias: defined as all AEs meeting the criteria for the SMQ "Dyskinesia" (narrow terms)
- Hallucinations and psychotic behaviors: defined as all AEs meeting the criteria for the SMQ "Psychosis and psychotic disorders" (narrow terms)
- Impulse control disorders: defined as all AEs meeting the following criteria.
    - HLT "Impulse control disorders"
    - HLT "Paraphilias 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 AEs with HLGT "Sleep disorders and disturbances" or the preferred term of "sudden onset of sleep"
- QT prolongation and ventricular arrhythmias: defined as all AEs meeting the criteria for the SMQ "Torsade de pointes /QT prolongation" (Broad terms)
- Acute Coronary Syndrome, Myocardial infarction, Angina: defined as all AEs meeting the criteria for the SMQ "Myocardial infarction" (Broad terms)
- Suicidal ideation & attempts: defined as all AEs meeting the criteria for the SMQ "Suicide/self-injury" (narrow terms)

- Melanoma: defined as all AEs meeting the criteria for the SMQ "Skin malignant tumors" (narrow terms)
- Stomatitis, Oral ulcers, Oral irritation: defined as all AEs meeting the criteria for the SMQ "Oropharyngeal disorders" (narrow terms)
- Allergic/sensitivity response to the formulation: defined as all AEs meeting the criteria for the SMQ "Hypersensitivity" (broad terms)

AESI will also be summarized by PT and actual dose level, i.e. last prescribed dose prior to AE.

The time to onset for the first occurrence of Part B AEs will be summarized by treatment sequence for the most common AEs (PTs or any AESI category reported in at least 10 subjects in either treatment sequence) for Part B safety population. Kaplan-Meier curves will be used to illustrate the time to first onset of AEs (days). For subjects without an event at study completion, date of last visit for that subject will be used as the censoring date. For any AESI category reported in at least 10 subjects in treatment sequence only the main category will be used, i.e. these summaries will not be broken down by PT or other sub-category. The first occurrence of a PT within the AESI (i.e. first PT meeting the criteria for the AESI) will be used for analysis and the Kaplan-Meier probability estimates will be summarized as time interval by weeks (1 week=7 days).

16.1.6. AEs LEADING TO DOSE INTERRUPTED/REDUCED

AEs leading to drug interruption of study treatment will be identified by using the AE action taken as "Drug Interrupted" from the AE page of the (e)CRF.

AEs leading to dose reduction of study treatment will be identified by using the AE action taken as "Dose Reduced" from the AE page of the (e)CRF.

For AEs leading to dose interruption and dose reduction of study treatment, summaries, both as event and subject counts by SOC and PT will be prepared, separately for Part A and Part B. Similar tables by actual dose level will also be summarized. The actual last dose level i.e. last prescribed dose prior to the AE onset will be used for this analysis.

16.2. LABORATORY EVALUATIONS

Clinical laboratory tests are performed at screening, V1, V4, EOS and ET visits. Results from the central laboratory will be included in the reporting of this study for Hematology, Chemistry and Urinalysis by overall.

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

X" in the listings.

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 for Part B Safety Population:

- Numeric laboratory parameters: Actual values and change from baseline (SV) to Week 0 (V1/V4) and EOS/ET, if applicable, for each parameter, will be summarized with descriptive statistics.
- Categorical laboratory parameters: The distribution of the categories will be summarized at SV, Week 0 (V1/ V4) and EOS/ET.

16.3. ECG EVALUATIONS

A standard 12-lead ECG will be performed at screening, TV1, APL DCVs, SC TV1 SC TV2, V1, V3, V4, V6, EOS, ET, and dose reduction visits during Part B. All ECGs will be obtained in the supine position, after the subject has been resting supine for at least 10 minutes. ECGs will be 12-lead with a 10-second rhythm strip. The ECG data will be centrally over-read and interpreted, and the following parameters will be captured: Heart rate, PR interval, QRS interval, RR interval, QT interval, QTcF Interval (Fridericia's correction) and QTcB Interval (Bazett's correction). A triplicate 12-lead ECG will be performed at screening and baseline, and the median value will be used for summary purposes. For Part A and Part B, ECGs will be performed just prior to dosing and 50 (± 30) minutes post-dose. Unscheduled ECGs may be performed if clinically indicated by changes in vital signs (orthostatic hypotension) or symptoms (new cardiovascular symptoms e.g., dizziness, syncope).

The following summaries will be done for PART A

- Actual values and changes from both baseline (last available screening value) and visit pre-dose to the post-dose timepoint for each parameter will be presented for APL: by dose level and then SC: by dose level. In case, more than one observation is reported at a dose level, the latest observed value will be used for calculation. However, all values will be presented in listing.
- The ECGs will be centrally read and deemed "Normal", "Abnormal, not clinically significant" and "Abnormal, clinically significant". These assessments will be tabulated for APL: by dose level and then SC: by dose levels.

The following summaries will be done for PART B:

- Actual values and changes from both baseline (last available screening value) and visit pre-dose to the post-dose time point, at each visit week 0 and 4 for each parameter (including the standing minus supine values), by treatment group

- The ECGs will be centrally read and deemed "Normal", "Abnormal, not clinically significant" and "Abnormal, clinically significant". These assessments will be tabulated by visit week 0 and 4 and time point.

16.3.1. ECG MARKEDLY ABNORMAL CRITERIA

The markedly abnormal analysis includes the tabulation of outlier ECG values using the criteria shown below. The outliers occurring at least once post-baseline during the study, including unscheduled visits will be summarized. In the summary, the highest post-baseline ECG value for a subject and within the treatment will be used.

The QTc Intervals fulfilling the following criteria will be tabulated separately by treatment group 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.

16.4. VITAL SIGNS

Vital sign measurements include heart rate (HR), respiratory rate (RR), blood pressure (BP), and body temperature. The BP values and heart rate will be assessed both in supine and standing position. Vital signs will be collected during

- Screening
- Part A: APL TV1, APL DCVs, SC TV1 SC TV2 visits, prior to dosing and 60 minutes post-dose at each visit Post-dose ECG may be used as next pre-dose ECG.
- Part B: V1, V2, V3, V4, V5, V6, EOS/ET and dose reduction visits, prior to dosing and 30 minutes post-dose at each visit (90 and 120 minutes post-dose may be collected if clinically necessary).

In addition to the vital signs captured on the Case Report Form, evaluation for orthostatic hypotension (OH) will be conducted, using the standing minus supine values (standing minus supine systolic BP, standing minus supine diastolic BP).
The following summaries will be done for PART A

- Actual values and changes from both baseline (last available screening value) and visit pre-dose to the post-dose time point: 60 minutes for Part A, for each parameter (including the

standing minus supine values), will be presented for APL-130277: by dose level and then SC: by dose level.

For those visits where more than one dose level is administered, if pre-dose for a dose level is missing, then use the pre-dose value for the first dose level administered on that visit day. E.g. On SC TV1, subject receives dose level 2 mg, 3 mg, and 4 mg. If pre-dose value for 4 mg is missing, then use the 2 mg pre-dose value as pre-dose for 4 mg. In case, more than one observation is reported at a dose level, the latest observed value will be used for calculation. However, all values will be presented in listing.

- Orthostatic hypotension will be defined as a reduction in systolic BP of 20 mmHg or more, and/or a reduction in diastolic BP of 10 mmHg or more, for the standing measurement compared to the supine measurement. The proportion of subjects with orthostatic hypotension will be presented for APL-130277: by dose levels and then SC: by dose levels

The following summaries will be done for PART B:

- Actual values and changes from both baseline (last available screening value) and visit pre-dose to the post-dose time point(s): 30 minutes for Part B, at each visit week for each parameter (including the standing minus supine values), by treatment group.
- Orthostatic hypotension will be defined as a reduction in systolic BP of 20 mmHg or more, and/or a reduction in diastolic BP of 10 mmHg or more, for the standing measurement compared to the supine measurement. The proportion of subjects with orthostatic hypotension will be tabulated by visit week and time point, by treatment group.

16.5. PHYSICAL EXAMINATION

Whether physical examination was performed or not will be listed. In addition, the frequency and percentage of subjects with each type of oropharyngeal cavity examination and injection site exam finding will be summarized by dose levels, time point (when applicable) and location. Clinically significant findings for physical examination, oropharyngeal cavity examination and injection site exam are recorded as AEs.

16.6. OTHER SAFETY ASSESSMENTS

16.6.1. COLUMBIA SUICIDE SEVERITY RATING SCALE (C-SSRS)

The Columbia Suicide Severity Rating Scale 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 subjects with each response for suicidal ideation, intensity of ideation, and suicidal behavior items will be summarized

as appropriate by treatment group and visit week for PART A and B. Clinically significant C-SSRS findings are recorded as AEs.

16.6.2. QUESTIONNAIRE FOR IMPULSIVE-COMPULSIVE DISORDERS IN PARKINSON'S DISEASE – RATING SCALE (QUIP-RS)

This is an instrument used to measure the extent of impulsive and compulsive behaviors in PD subjects. The QUIP-RS consists of four questions which are 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. Hobbyism and punding are combined to get a score range of 0-32. The frequency and percentage of subjects with positive response ("rarely" or higher) for each disorder will be summarized by treatment group and visit week for the Titration phase and the Maintenance phase. 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 week (both absolute values and changes from baseline (V1)) for PART B.

17. REFERENCES

Ratitch, B. and O'Kelly, M. (2011), "*Implementation of Pattern-Mixture Models Using Standard SAS/STAT Procedures*," in Proceedings of PharmaSUG 2011 (Pharmaceutical Industry SAS Users Group), SP04, Nashville.

Dobson, A. J. (2002). *An introduction to generalized linear models, 2nd ed.* Chapman & Hall/CRC.

EMA. (n.d.). *Guideline on Missing Data in Confirmatory Clinical Trials.* Retrieved December 15, 2011, from European Medicines Agency: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500096793.pdf

Little, R., & Yau, L. (1996). Intent-to-Treat Analysis for Longitudinal Studies with Drop-Outs. *Biometrics* , vol 52, 1324-1333.

Tao J, Kiernan K, Gibbs P. Advanced Techniques for Fitting Mixed Models Using SAS/STAT® Software. SAS Institute white paper number SAS1919; 2015.

Yang Yuan (2014). *Sensitivity Analysis in Multiple Imputation for Missing Data*, Paper SAS270-2014.