



Protocol B7601003

**A 15-WEEK, PHASE 2, DOUBLE BLIND, RANDOMIZED,
PLACEBO-CONTROLLED, DOSE RANGING STUDY TO INVESTIGATE THE
EFFICACY, SAFETY AND TOLERABILITY OF PF-06649751 IN SUBJECTS WITH
MOTOR FLUCTUATIONS DUE TO PARKINSON'S DISEASE**

**Statistical Analysis Plan
(SAP)**

Version: 3

Date: 13 September 2017

TABLE OF CONTENTS

LIST OF TABLES5

LIST OF FIGURES5

APPENDICES5

1. VERSION HISTORY6

2. INTRODUCTION9

 2.1. Study Objectives10

 2.1.1. Primary Objective10

 2.1.2. Secondary Objectives10

 2.2. Study Design11

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS13

 3.1. Primary Endpoint(s)13

 3.2. Secondary Endpoint(s)17

 3.2.1. MDS-UPDRS Part III18

 3.2.2. MDS-UPDRS Parts I, II, IV, and Total Score19

CCI [Redacted]

CCI [Redacted]

090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

CCI

3.4. Baseline Variables	26
3.5. Safety Endpoints	27
3.5.1. Adverse Events	27
3.5.2. Laboratory Data	27
3.5.3. Vital Signs (Blood Pressure and Pulse Rate).....	28
3.5.4. Electrocardiogram (ECG).....	28
3.5.5. Physical Examination	29
3.5.6. Neurological Examination	29
3.5.7. Assessment of Suicidal Ideation and Behavior	29
3.5.7.1. Columbia Suicide Severity Rating Scale (C-SSRS)	29
3.5.7.2. Mental Health Risk Assessment (MHRA)	32
3.5.8. Questionnaire for Impulsive-Compulsive Disorders in Parkinson’s Disease – Rating Scale (QUIP-RS).....	32
3.5.9. Physician Withdrawal Checklist (PWC-20)	33
4. ANALYSIS SETS	33
4.1. Full Analysis Set	34
4.2. Per Protocol Analysis Set.....	34
4.3. Safety Analysis Set.....	34

CCI

5. GENERAL METHODOLOGY AND CONVENTIONS.....	35
5.1. Hypotheses and Decision Rules	35
5.1.1. Statistical Hypotheses.....	35
5.1.2. Statistical Decision Rules	35
5.2. General Methods	36
5.2.1. Analyses for Binary Data.....	37
5.2.2. Analyses for Continuous Data	37
5.2.2.1. Mixed Model for Repeated Measures Analysis (Non- Bayesian Approach)	37
5.2.2.2. Dose Response Analysis (Non-Bayesian Approach)	38

090177e18f6e873aApprovedOn: 26-Sep-2018 05:14 (GMT)

5.2.2.3. Bayesian Analysis	39
5.2.2.4. Dose Response Trend Test	41
5.2.3. Analyses for Categorical Data	41
5.3. Methods to Manage Missing Data	41
6. ANALYSES AND SUMMARIES	42
6.1. Primary Endpoint(s)	42
6.1.1. Change from Baseline in Daily OFF time at Week 10 (end of Period A)	42
6.1.1.1. Primary Analysis	42
6.1.1.2. Sensitivity/Robustness Analyses	43
6.2. Secondary Endpoint(s)	44
6.2.1. Hauser Home Diary	44
6.2.2. MDS-UPDRS	44
CCI	
6.4. Subset Analyses	47
6.5. Baseline and Other Summaries and Analyses	47
6.5.1. Study Conduct and Subject Disposition	47
6.5.2. Baseline Summaries	48
6.5.3. Study Treatment Exposure	48
6.5.4. Concomitant Medications and Non-Drug Treatments	49
6.6. Safety Summaries and Analyses	49
6.6.1. Adverse Events	49
6.6.2. Laboratory Data	50
6.6.3. Vital Signs (Blood Pressure and Pulse Rate)	50
6.6.4. Electrocardiogram	51
6.6.5. Physical Examination	52
6.6.6. Neurological Examination	52
6.6.7. Assessment of Suicidal Ideation and Behavior	52
6.6.7.1. Columbia Suicide Severity Rating Scale (C-SSRS)	52
6.6.7.2. Mental Health Risk Assessment (MHRA)	52

6.6.8. Questionnaire for Impulsive-Compulsive Disorders in Parkinson’s Disease – Rating Scale (QUIP-RS).....	53
6.6.9. Physician Withdrawal Checklist (PWC-20)	53
7. INTERIM ANALYSES	53
7.1. Introduction	53
7.2. Interim Analyses and Summaries.....	54
8. REFERENCES	55
9. APPENDICES	56

LIST OF TABLES

Table 1. Summary of Major Changes in SAP Amendments	6
Table 2. <i>Proposed Titration Schemes for PF-06649751 and Placebo Doses</i>	12
Table 4. C-SSRS Mapped to C-CASA - Suicidality Events and Codes.....	30
Table 4. C-SSRS Mapped to C-CASA – Examples of Worsening/New Onset	31
Table 5. Safety QTcF ¹	51

LIST OF FIGURES

Figure 1. Study Schematic	12
---------------------------------	----

APPENDICES

Appendix 1. SUMMARY OF EFFICACY ANALYSES.....	56
Appendix 2. DATA SET DESCRIPTIONS	59

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study B7601003 Version 3 is based on the protocol amendment 2 (A2) dated 22NOV2016 (approval date 14Dec16).

Table 1. Summary of Major Changes in SAP Amendments

SAP Version	Change	Rationale
1	Not Applicable	Initial version
2	Section 2	<ul style="list-style-type: none"> CCI Toxicology data updated Included text on the B7601017 extension study
2	CCI	
2	Section 2.2 and Figure 1	Days 14 & 21 may be phone visits
2	Section 2.2	Added text in regards to Japan specific visits.
2	Sections 2.2 and 7.1	Interim Analysis language clarified.
2	Section 3.1	Updated Hauser diary rules:

090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

		<ul style="list-style-type: none"> Accounted for Pre-A2 diaries collected anytime during screening period Addressed impact if Week 3 (Day 21) is either an in-clinic or phone visit Modified rescue medication methods to include Period B for subjects Post A2 Noted Week 13 is only relevant Pre-A2 Removed rules for subjects with more than one observation for the same 30 minute time period. Data will be considered invalid and set to missing.
2	Sections 3.2, 6.2.1 and Appendix 1	“Daily ON time with troublesome dyskinesia (hours)” added as a secondary endpoint
2	Section 3.4	Baseline is noted as Day 0 throughout, which is consistent with the terminology in the protocol. However, for purposes of reporting it will be referred to as Day -1 which is consistent with CDISC standards.
2	Section 3.2.1	Added lip/jaw to the MDS-UPDRS Part III tremor subscale and updated rule for missing items.
2	CCI	
2		
2		
2		
2	Sections 3.5.3 and 6.5.2	Added temperature.

2	Section 3.5.7 and 6.6.7	<ul style="list-style-type: none"> • C-SSRS “Lifetime” replaced with “Lifetime/Past 12 months” throughout. • Moved C-SSRS as a subsection of Assessment of Suicidal ideation and Behavior, and added Mental Health Risk Assessment (MHRA) subsections.
2	Section 4.2	Updated Per Protocol Analysis Set (PPAS) criteria
2	Section 5.2	All available visit data will be reported, and for descriptive statistics by time point each visit will be represented.
2	Section 5.3	Added general clarification on missing data at Baseline to use Screening values as Baseline (Day 0).
2	Section 5.3	Added handling of the early termination visit for safety.
2	Section 6	Clarified that data after Week 10 Pre-A2 will be in listings only, otherwise all data will be presented in summaries and used in statistical modelling.
2	Section 6.6.5	Removed non-standard table, since clinically significant abnormalities which meet the definition of an adverse event will be recorded as such.
2	Section 6.6.6	Neurological examination data will be presented now as only listings, since all newly abnormal findings are required to be logged as AEs.
3	Sections 3.1 and 6	Clarify impact of Protocol Amendment 2. Administrative change to align with the IA #1 SAP.
3	Section 3.1	Added a rule to address excluding home diaries that are not within 7 days of the clinic visit. If the excluded home diary nominal visit is equal to the closet planned visit it will now be included, otherwise it is still excluded.

3	Section 3.5.1	Adapted the “Rule of 4” for Tier 2 events.
3	Section 3.5.8	QUIP-RS was updated as follows: <ul style="list-style-type: none"> • Now 8 scoring types • Updated missing responses rules
3	Section 6.4	Added subset analysis based on L-Dopa intake during baseline. Explicitly stated for all subset analyses, a minimum number of subjects required for a subgroup to be analyzed.
3	Section 6.6.8	Added additional details on summarizing the QUIP-RS.
3	Appendix 1	Added that subset analyses were optional, as they require $\geq 10\%$ of total subjects to be analyzed.

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B7601003. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

Text taken directly from the protocol is italicized.

PF-06649751 is a highly selective partial agonist at dopamine D1 like receptors (D1 and D5 receptors, abbreviated as DIRs) which is being developed for the treatment of the signs and symptoms of Parkinson’s disease. In Parkinson’s disease, motor function can be pharmacologically rescued with activation of either or both DIR-containing direct and D2R-containing indirect striatal output pathways. In contrast to available D2/D3R agonists, D1/D5R agonists have demonstrated efficacy similar to L-Dopa in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) lesioned nonhuman primate model of Parkinson’s disease (see Investigators Brochure Section 5.1.3.2, In Vivo Pharmacodynamics). Severely lesioned MPTP-treated monkeys showed no response to D2 agonists and modest improvement with L-Dopa treatment, but showed marked improvement with D1 agonist treatment.

PF-06649751 is a novel D1 and D5 specific dopamine partial agonist and has the potential to reduce OFF time in subjects with moderate to advanced Parkinson’s disease who are experiencing significant OFF time despite their current dopaminergic therapy. Following an

initial titration phase, which is intended to mitigate potential dopaminergic adverse events such as nausea and vomiting, fixed doses of PF-06649751 will be evaluated as adjunctive treatment with L-Dopa. The overall doses of PF-06649751 selected for this study were influenced by the aim to identify not only a minimally-effective dose but also characterize the maximum efficacious dose. PF-06649751 doses for this study are selected based on evidence of pharmacological activity from preclinical data and clinical studies in Parkinson's disease subjects (See the Protocol for more detail).

The study was initially limited to a maximum 15 week treatment period due to the available toxicological coverage. Subsequently, the nonclinical safety profile of PF-06649751 is considered to be adequately characterized to support progression into human clinical studies of 6 months or longer dosing duration for men and for women of non-childbearing potential.

Subjects who complete through Week 15 of the B7601003 study will have the option to enroll in an open label extension study (B7601017). If the pre-specified Inclusion/Exclusion criteria are met at Week 10 (screening), the subject may exit the current B7601003 study at Week 15 for continued treatment and follow-up under the extension study (B7601017), provided they remain eligible. Subjects not directly entering study B7601017 at Week 15 must complete the Week 17 and Week 19 Follow-up visits and must rescreen under the B7601017 protocol.

The study population will include male subjects and female subjects of non-childbearing potential diagnosed with Parkinson's disease who experience motor fluctuations.

This study will evaluate the efficacy, safety, and tolerability of up to 4 different doses of PF-06649751 in subjects with Parkinson's disease as adjunctive therapy to L-Dopa.

2.1. Study Objectives

2.1.1. Primary Objective

- *To evaluate the effect on motor symptoms of PF-06649751 administered once daily as adjunctive treatment with stable doses of L-Dopa in Parkinson's disease.*
- *To determine the therapeutic window for motor symptom improvement of PF-06649751 administered once daily, ie, determining a dose, or a range of doses, for adequate control of motor symptoms.*

2.1.2. Secondary Objectives

- *To evaluate the dose-response on motor symptoms of PF-06649751 administered once daily in subjects with Parkinson's disease.*
- *To evaluate the safety and tolerability of PF-06649751 administered once daily in subjects with Parkinson's disease.*

- *To evaluate the effect of PF-06649751 administered once daily on dyskinesia in subjects with Parkinson's disease.*

CCI

2.2. Study Design

The study has a randomized, double-blind, placebo-controlled parallel group design. Approximately 198 subjects from approximately 60 centers in approximately 6 countries will be randomized to 5 treatment groups (15 mg QD, 7 mg QD, 3 mg QD, 1 mg QD, or placebo), initially in a 2:0:0:0:1 ratio of approximately 54 subjects (Cohort 1) followed by a 1:2:2:2:1 ratio of approximately 144 subjects (Cohort 2), using a central randomization system with randomization blocks stratified by region. Geographic region will be defined as North America, Japan, and Europe.

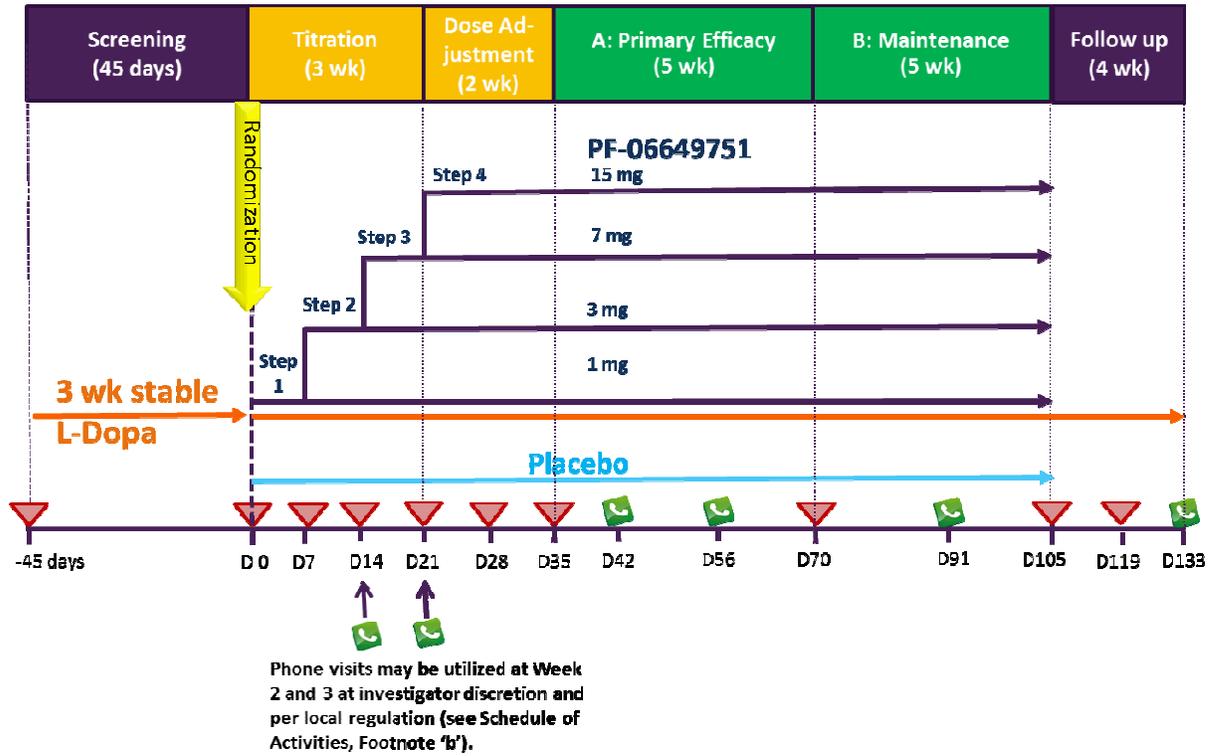
Each subject will participate in the study for approximately 25 weeks including up to a 45 day screening period, 15 week double-blind treatment period (comprising up to a 3 week titration, 2 week dose adjustment, 5 weeks of primary efficacy with adjunctive treatment to stable L-Dopa (Period A), and 5 weeks of maintenance with adjunctive treatment to stable L-Dopa (Period B)), and an approximately 28 day follow-up period.

Note that prior to B7601003 Protocol Amendment 2 Days 14 and 21 were required in-clinic visits, but now with the option to be phone visits, as applicable. The impact of this change will be noted in subsequent sections.

For sites in Japan, the Week 2 (Day 14) and Week 3 (Day 21) visits during the Titration Period must be conducted as site (in-clinic) visits. Additionally, two additional phone visits are required a day after both the Day 0 (Randomization) Visit (Day 1) and Week 1 Visit (Day 7 [+3 day window]) and two additional study site visits are required a day after both Week 2 Visit (Day 14 [+3 day window]) and Week 3 Visit (21 [+3 day window]), ie Day 15 and Day 22 respectfully. Phone visits can optionally be conducted as site visits (eg, for logistical reasons).

The study design is illustrated in the figure below.

Figure 1. Study Schematic



Each subject is planned to be up-titrated in a double-blind fashion to the assigned target dose level of PF-06649751 or placebo according to the following titration scheme:

Table 2. Proposed Titration Schemes for PF-06649751 and Placebo Doses

Week/ Start Day of Dose Level	Treatment Group				
	Placebo	1 mg	3 mg	7 mg	15 mg
Week 1 / from Day 1	P	1 mg	1 mg	1 mg	1 mg
Week 2/ from Day 8	P	1 mg	3 mg	3 mg	3 mg
Week 3 / from Day 15	P	1 mg	3 mg	7 mg	7 mg
Weeks 4-15 / from Day 22	P	1 mg	3 mg	7 mg	15 mg

Titration (Week 1 to Week 3): *No modifications to the upward titration of Study Medication can be made. Subjects who experience intolerable adverse events during titration will need to be discontinued from the study.*

Post Week 3 (Week 4 to Week 5): *For subjects who experience unacceptable new adverse events after having achieved the target dose (after Day 21) and deemed related to the dose increase, the dose of investigational product can be changed to the previous dose level (Week 3 dose level). This results in a dose reduction from 15 mg/day to 7 mg/day for the highest dose group, and no change in dose for the other dose groups in a blinded fashion. Only one dose modification of investigational product is permitted.*

L-Dopa Adjustment for Unacceptable Dopaminergic Side Effects (Week 1 to Week 15): *In the case of unacceptable dopaminergic side effects and following discussion with the Medical Monitor, the dose of L-Dopa may be reduced between Day 1 and Day 105 visit.*

Period A: Primary Efficacy: *During Period A, subjects will be administered double blind study medication once daily (up to 4 different dose levels of investigational product) as adjunctive treatment to a stable dose of open-label L-Dopa.*

Period B: Maintenance: *After the Week 10 Visit (primary endpoint collection), subjects will enter Period B. During Period B, double-blind treatment with investigational product will continue as adjunctive treatment to a stable dose of open label L-Dopa. Note that prior to B7601003 Protocol Amendment 2 Period B allowed the open label L-Dopa dose to be reduced following a recommended downward titration based on the L-Dopa dose at the end of Period A. The impact of this change will be noted in subsequent sections.*

Up to two interim analyses may be performed when approximately 27% and 80% subjects are randomized and have the opportunity to complete through Period A of the study. The first interim analysis will be conducted when Cohort 1 (approximately 54 subjects in a 2:1 ratio of 15 mg QD and placebo groups respectively) is completed through Period A and efficacy data have been collected, databased and cleaned. A second interim analysis may be undertaken when approximately 104 subjects complete through Period A of Cohort 2 (1:2:2:2:1 in 15 mg QD, 7 mg QD, 3 mg QD, 1 mg QD, and placebo respectively) and efficacy data have been collected, databased and cleaned. The purpose of these interim analyses will be to assess study futility and to aid future development planning. If an interim analysis is performed, the study will continue whilst the analysis is being conducted.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoint(s)

The primary endpoint is the change from baseline in daily OFF time at Week 10 (end of Period A). The daily OFF time is calculated as the average of the three consecutive daily OFF hours from the symptom diary at each visit.

A paper Hauser diary, otherwise referred to as the “Parkinson’s Disease Home Diary”, will be utilized to record patient motor state for half hour intervals. Entries are made on the days noted in the STUDY PROCEDURES section of the protocol. Training on the diary will be provided to sites and subjects. Subjects will complete the diary for three consecutive days in the week prior to each visit (except Day 28 visit) as indicated in the STUDY PROCEDURES section, including three consecutive days during the week prior to Day 0 (Randomization). Note that prior to B7601003 Protocol Amendment 2, the three consecutive days prior to Day 0 could occur anytime during the 45 day screening period. Rules will be implemented in this section to ensure that data is retained for analysis.

Subjects are asked to make an **independent** assessment of motor fluctuations; caregivers or site study staff cannot participate (caregivers are not to provide opinions or tell the subject what he/she has observed). The subject will be instructed to answer the PD home diary on whether the subject has been ON with no dyskinesias, ON with non-troublesome dyskinesia, ON with troublesome dyskinesia, OFF, or ASLEEP. The home diary is completed for the full 24 hours reflecting both time awake and time asleep. A home diary day starts with the interval 24:00-0:30 thru 23:30-24:00 on each chronological day for 3 consecutive days. On the days recording the home diary, subjects will be instructed to make an entry every 30 minutes during their normal waking time and upon awakening from time asleep. Double entries and unclear entries (e.g. participant checks in the middle of 2 options) are considered invalid and entered as missing data, which is handled at the point of data entry via these established rules.

An average daily OFF time (using 3 days recording the Hauser diary prior to the visit) will be calculated at all time points for each subject, which includes Baseline (prior to Day 0), Week 3, Week 5 and Week 10 (End of Period A), Week 13 (Period B; exists prior to Protocol Amendment 2 only) and also Week 15 (End of Period B). Change from baseline will then be calculated. Week 10 is the primary endpoint.

- Week 3 Hauser home diary data will be collected at the Week 3 visit for subjects randomized prior to Protocol Amendment 2, and at the Week 3 visit for subjects randomized under Protocol Amendment 2, where sites choose to perform an in-clinic visit (instead of a phone visit).
- Week 3 Hauser home diary data will be collected at Week 4 for subjects randomized under Protocol Amendment 2, when sites perform a phone visit for Week 3. The Week 3 phone visit will be used as the nominal visit.

In calculating the average daily OFF time, the following algorithms and rules are applied. (Rescue medication methods below are for all periods through Period A, and additionally Period B for subjects consented under Protocol Amendment 2 prior to entering Period B).

Valid Hauser Diary Days

1. Subjects who do not complete at least 12 awake hours of symptom diary data (not including ASLEEP time) will have that daily data set to missing (i.e. will not be considered as a valid symptom diary day). There will be no imputation of data for these diary days.
2. Any Hauser diary day on which rescue medication was used will NOT be considered a valid symptom diary day. The Hauser diary data reported on that day hereafter referred to as a rescue diary day, will not be used. As described more completely in the instruction for daily OFF time below, baseline data will be used to impute the observed response on these days. (All periods through Period A, and additionally Period B for subjects consented under Protocol Amendment 2 prior to entering Period B).

Entries at 30-minute Intervals (Only for subjects who have ≥ 12 awake hours of Hauser home diary data)

3. If a single entry of 30 minutes is missing (ie, adjacent time points to that missing entry are not missing), an imputation algorithm will be employed using the adjacent available diary values carried forward and backward. If the missing time period was the first half-hour of the diary, the value is carried back from the second time period. If the missing time period was the last half-hour of the diary, the value is carried forward from the second to last time period. For missing time periods between the first and last half-hours, the first 15 minutes is imputed by carrying forward the value(s) from the previous half-hour and the second 15 minutes is imputed by carrying back the value(s) from the succeeding half-hour. This imputation will be completed prior to proceeding to the next steps, and will include ASLEEP time.
4. If data are missing for two or more consecutive 30-minute intervals between recorded ASLEEP times, fill in the missing 30-minute intervals with ASLEEP. Otherwise, record any missing between 10 PM and 6 AM as ASLEEP.
5. If data is missing for two or more consecutive 30-minute intervals, an imputation algorithm will be employed for each 30-minute interval using the available data at the other time intervals across the whole Hauser diary day.
 - Determine the % of awake proportions (no rounding), using available data for that home diary day, for ON with no dyskinesias, ON with non-troublesome dyskinesia, ON with troublesome dyskinesia, and OFF time. For example (rounding done for simplicity), 14 waking hours, 8 hours ON with no dyskinesias (57.14%), 4 hours ON with non-troublesome dyskinesia (28.57%), 1 hour ON with troublesome dyskinesia (7.14%), and 1 hour OFF time (7.14%).

- Multiply each of these proportions by 30 (round to integer) to have the imputation minutes for each of the missing 30-minute diary intervals. For example, 17 minutes (30×0.5714) ON with no dyskinesias, 9 minutes ON with non-troublesome dyskinesia (28.57%), 2 minutes ON with troublesome dyskinesia (7.14%), and 2 minutes OFF time (7.14%).
- Each missing 30 minute interval under this imputation will be similarly imputed with the minutes in the above calculation.

Daily OFF Time

6. All Hauser home diary data will be determined for a 24-hour period. All times are expressed in decimal hours (two decimal places).
7. The daily OFF time for every rescue diary day will be replaced by the subject's baseline OFF time, that is, the baseline 3-day average. (All periods through Period A, and additionally Period B for subjects consented under Protocol Amendment 2 prior to entering Period B).

Daily Average (Rescue diary days are for all periods through Period A, and additionally Period B for subjects consented under Protocol Amendment 2 prior to entering Period B)

8. Valid home diary days (at least 12 awake hours, after all imputation) and/or rescue diary days will be used to compute the average daily OFF time. The valid home or rescue diary day must be within 7 days of the clinic visit (or phone visit for Week 3 only as noted below or assigned to the closest planned visit as noted below) but not on or after the day of the clinic visit. If more than 3 valid home or rescue diary days are available, the 3 days closest to the clinic visit will be used.
 - Prior to Protocol Amendment 2, the three consecutive days before Day 0 could occur anytime during the screening period and was not restricted to within a week of the visit. Therefore, all home diaries before Day 0 will be considered for Baseline, with the 3 days closest to Day 0 being used for analysis.
 - Subjects, who have a phone visit for Week 3 and their home diary collected at Week 4, will use their Week 3 phone visit to determine within 7 days.
 - If a valid home or rescue diary is considered for exclusion based on the "within 7 days" rule, it will still be included if its nominal visit is equal to the closest planned visit, otherwise it will be excluded. Planned visits: Baseline (Day -1), Week 3 (Day 21), Week 5 (Day 35) or Week 10 (Day 70).
 - i. Example 1: A subject who completed Week 3 diaries on Days 8, 9, and 10 with the actual Week 3 visit occurring on Day 21. All diaries excluded as all Week 3 diaries would all have been closer to Baseline (Day -1) than Week 3 (Day 21).

- ii. Example 2: A subject who completed Week 10 diaries on Days 61, 62 and 63 with the actual Week 10 visit occurring on Day 70. All diaries included because the diaries are closest to Day 70 which matches the clinic visit (Week 10).
 - iii. Example 3: A subject who completed their Week 3 diaries on Days 11, 12 and 13 but the actual Day 21 visit occurred on Day 24. All diaries included because the diaries are closest to Day 21 which matches the clinic visit (Week 3).
9. If only 2 valid home or rescue diary days are available prior to a visit, data from the 2 days will be used to calculate the average daily OFF time.
 10. If only one valid home or rescue diary day is available the average daily from the last available week will be averaged with the daily OFF time from the one valid diary day.
 11. Subjects that do not have any valid home diary days for a visit or who are completely missing a visit will have the average daily OFF time set to missing for that visit.

The extent of missing data will be evaluated as part of the Blind Data Review Meeting(s). Tabular displays will be developed to describe the extent of missing data.

3.2. Secondary Endpoint(s)

Secondary endpoints evaluated during the entire double-blind period include:

- Daily OFF time (hours): change from baseline in average daily OFF time by Visit.
- Daily ON time with troublesome dyskinesia (hours): change from baseline in average daily ON time with troublesome dyskinesia by Visit.
- Daily ON time without troublesome dyskinesia (hours): the change from baseline in average daily ON time without troublesome dyskinesia (i.e. without dyskinesia or with non-troublesome dyskinesia, or alternatively “good time”) by Visit.

The same algorithm and rules employed for the calculation of OFF time will be used for the secondary variables. Baseline is defined using the 3 Hauser diary days prior to Day 0.

- MDS – Unified Parkinson’s Disease Rating Scale (MDS-UPDRS) Part III: Change from Baseline in MDS-UPDRS Part III Score by Visit.
- MDS-UPDRS Parts I, II, IV, and total score: Change from Baseline in the MDS-UPDRS Parts I, II, IV and Total Scores by visit.

The MDS-UPDRS will be collected at times specified in the STUDY PROCEDURES section of this protocol. *Although not required, every effort should be made to have the same individual perform the ratings for an individual subject throughout the course of the study.*

MDS-UPDRS Part III should be assessed consistently and approximately 2 to 4 hours following L-Dopa dose (except at screening).

3.2.1. MDS-UPDRS Part III

MDS-UPDRS Part III assesses the motor signs of Parkinson's disease and is administered by the investigator. It is comprised of 33 sub-scores based on 18 items, several with right, left or other body distribution scores. Each question is anchored with five responses that are linked to commonly accepted clinical terms: 0 = normal, 1 = slight, 2 = mild, 3 = moderate, and 4 = severe. Assessments will be done at times specified in the STUDY PROCEDURES section of the protocol. Higher total scores indicate more severe motor signs of Parkinson's disease. There are four subscales:

- The tremor subscale, which is the sum of scores for resting tremor amplitude in five body parts (5 items: right upper extremities, left upper extremities, right lower extremities, left lower extremities, and lip/jaw), postural tremor (2 items) and Kinetic tremor (2 items) in the hands (range 0–36).

If more than 1 item (>20%) is missing for a time point, the subscale score for that time point is missing; otherwise, the subscale score will be imputed as follows: sum of the non-missing items X (total number of items)/ (number of items non-missing).

- The rigidity subscale is the sum of rigidity scores for limbs and neck (5 items, range 0–20).

The subscale score will be missing if more than 1 item (>20%) is missing.

- The bradykinesia subscale is the sum of scores for finger tapping (2 items), hand movements (2 items), pronation and supination movements (2 items), toe tapping (2 items) and Global spontaneity of movement (body bradykinesia) (1 item) (range 0–36).

If more than 1 item (>20%) is missing for a time point, the subscale score for that time point is missing; otherwise, the subscale score will be imputed as follows: sum of the non-missing items X (total number of items)/ (number of items non-missing).

- The PIGD subscale is the sum of gait (1 item), freezing of gait (1 item) and postural stability (1 item) (range 0-12).

The subscale score will be missing if more than 1 item (>20%) is missing.

If more than 7 of the MDS-UPDRS Part III items are missing for a time point,⁵ the part score for that time point is missing; otherwise, the part score will be imputed as follows: sum of the non-missing items X (total number of items)/ (number of items non-missing).

Baseline will be the Day 0 measurement.

3.2.2. MDS-UPDRS Parts I, II, IV, and Total Score

MDS-UPDRS Parts I, II, and IV will be done at times specified in the STUDY PROCEDURES section of the protocol. Each question is anchored with five responses that are linked to commonly accepted clinical terms: 0 = normal, 1 = slight, 2 = mild, 3 = moderate, and 4 = severe. Higher part and total scores indicate more severe signs of Parkinson's disease.

- Part I: Non-Motor Aspects of Experiences of daily Living (nM-EDL); this portion of the scale assesses the non-motor impact of Parkinson's disease on patients' experiences of daily living. There are 13 questions. Part 1A is administered by the rater (six questions) and focuses on complex behaviors. Part 1B is a component of the self-administered Patient Questionnaire that covers seven questions on non-motor experiences of daily living. (range 0-52)

If more than 1 item is missing for a time point,⁵ the part score for that time point is missing; otherwise, the part score will be imputed as follows: $\text{sum of the non-missing items} \times (\text{total number of items}) / (\text{number of items non-missing})$.

- Part II: Motor Aspects of Experiences of Daily Living (M-EDL); this portion of the scale assesses the motor impact of Parkinson's disease on patients' experiences of daily living. There are 13 questions. (range 0-52)

If more than 2 items are missing for a time point,⁵ the part score for that time point is missing; otherwise, the part score will be imputed as follows: $\text{sum of the non-missing items} \times (\text{total number of items}) / (\text{number of items non-missing})$.

- Part IV: Motor Complications; in this section, the rater uses historical and objective information to assess two motor complications, dyskinesias and motor fluctuations that include OFF-state dystonia. Use all information from patient, caregiver, and the examination to answer the six questions that summarize function over the past week including today. (range 0-24)

The part score will be missing if any underlying question is missing.⁵

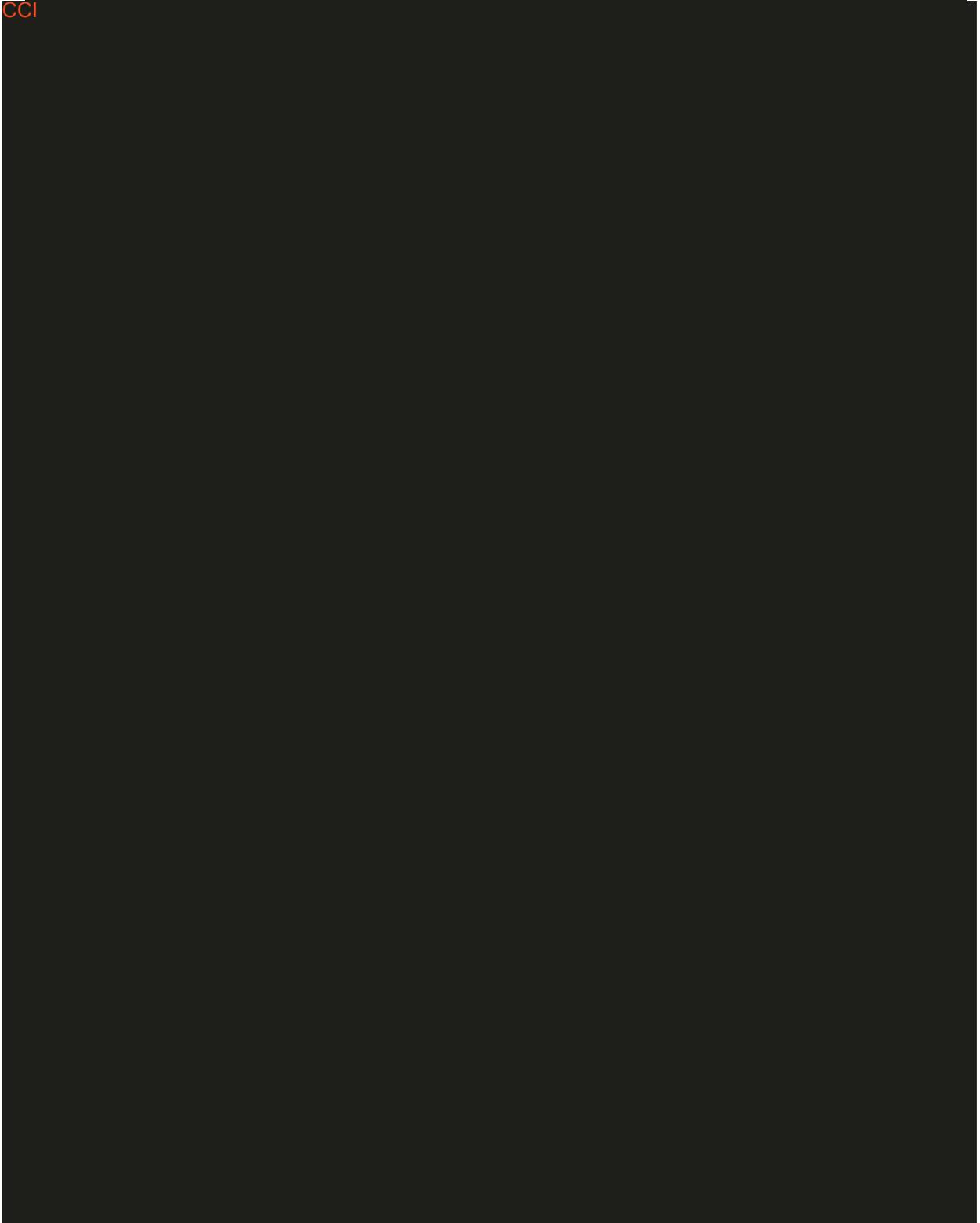
- MDS-UPDRS Total Score: The sum of Parts I, II, III, and IV.

The total scores will be missing if any underlying part score is missing.

Baseline will be the Day 0 measurement.

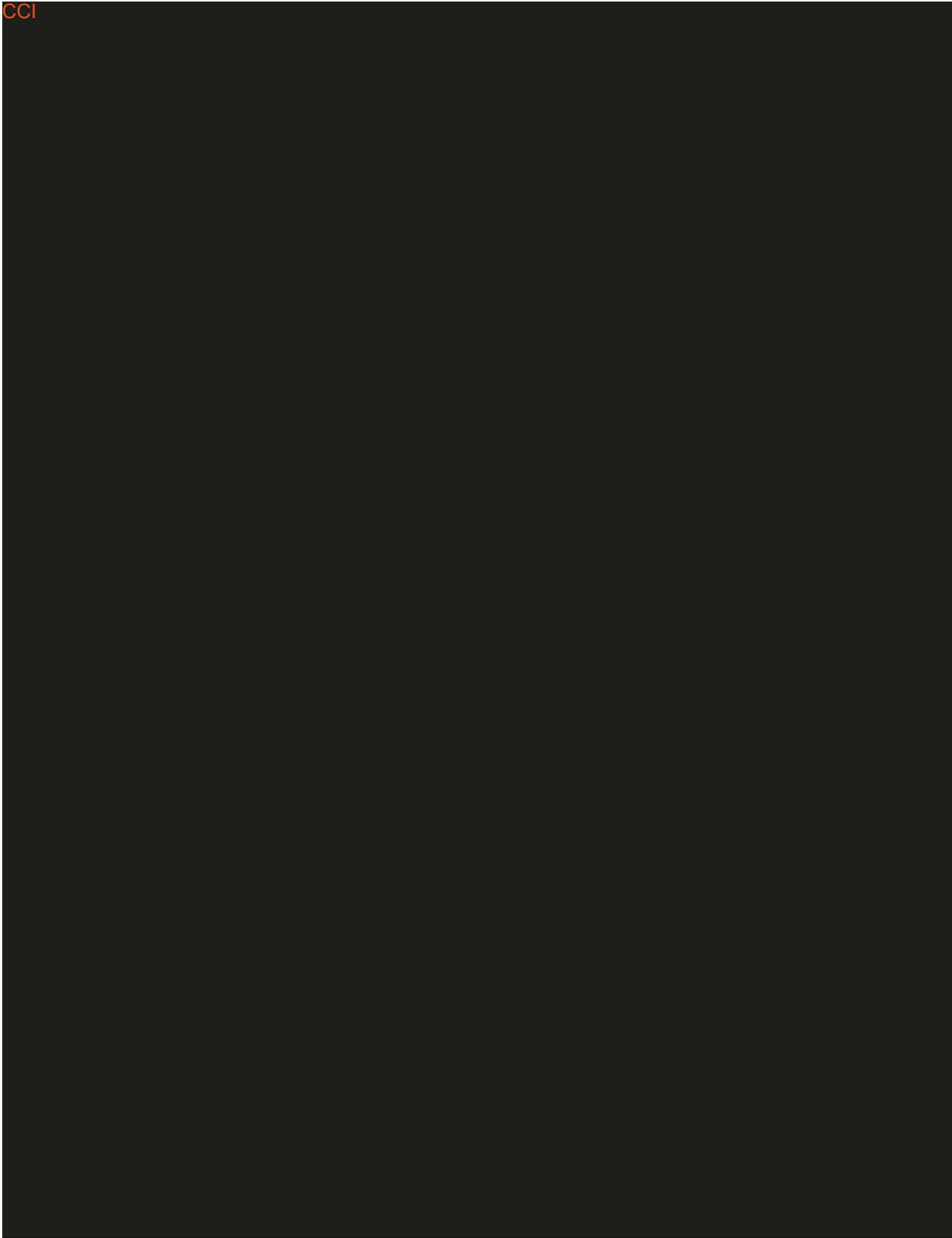
CCI

CCI



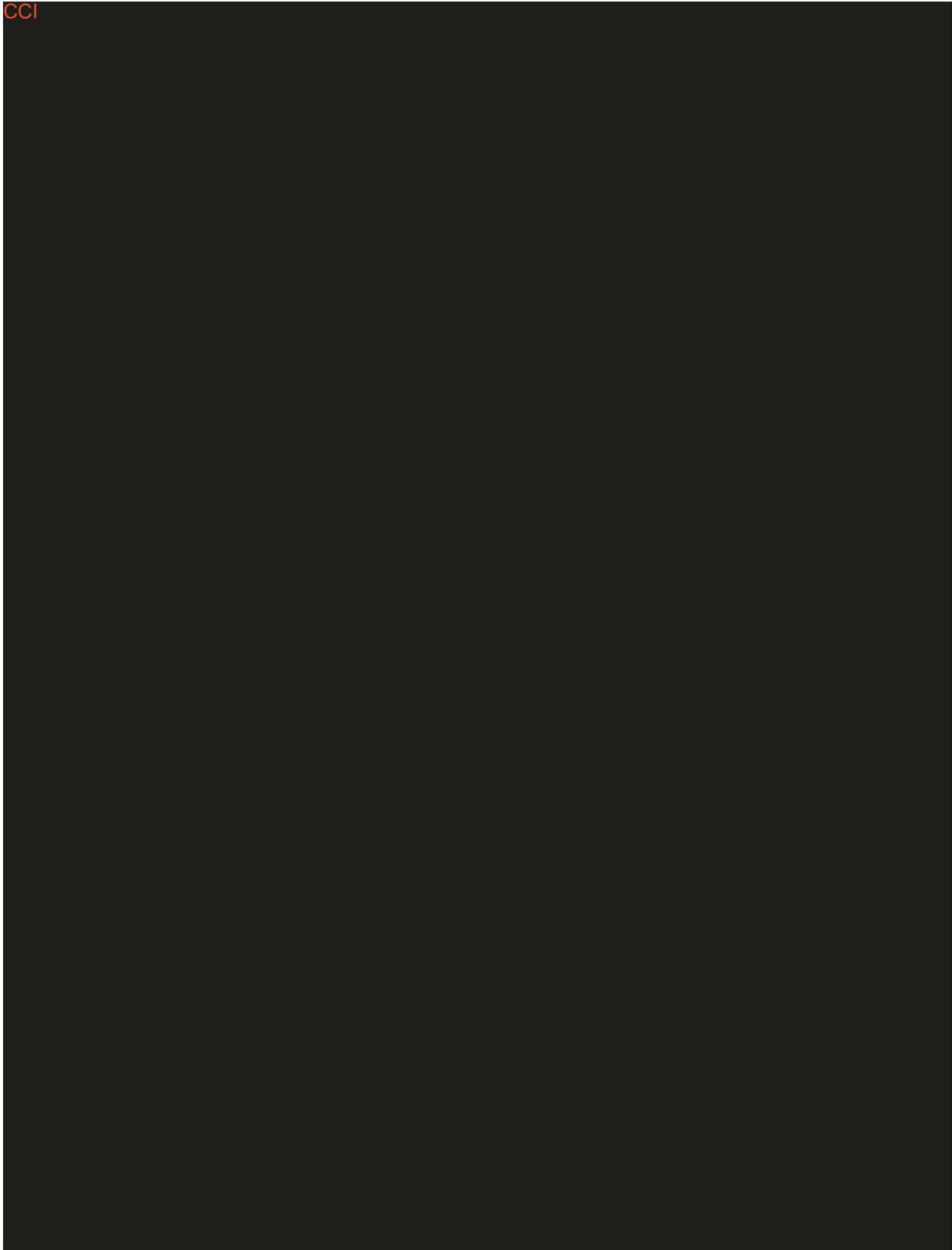
090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

CCI



090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

CCI



090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

CCI



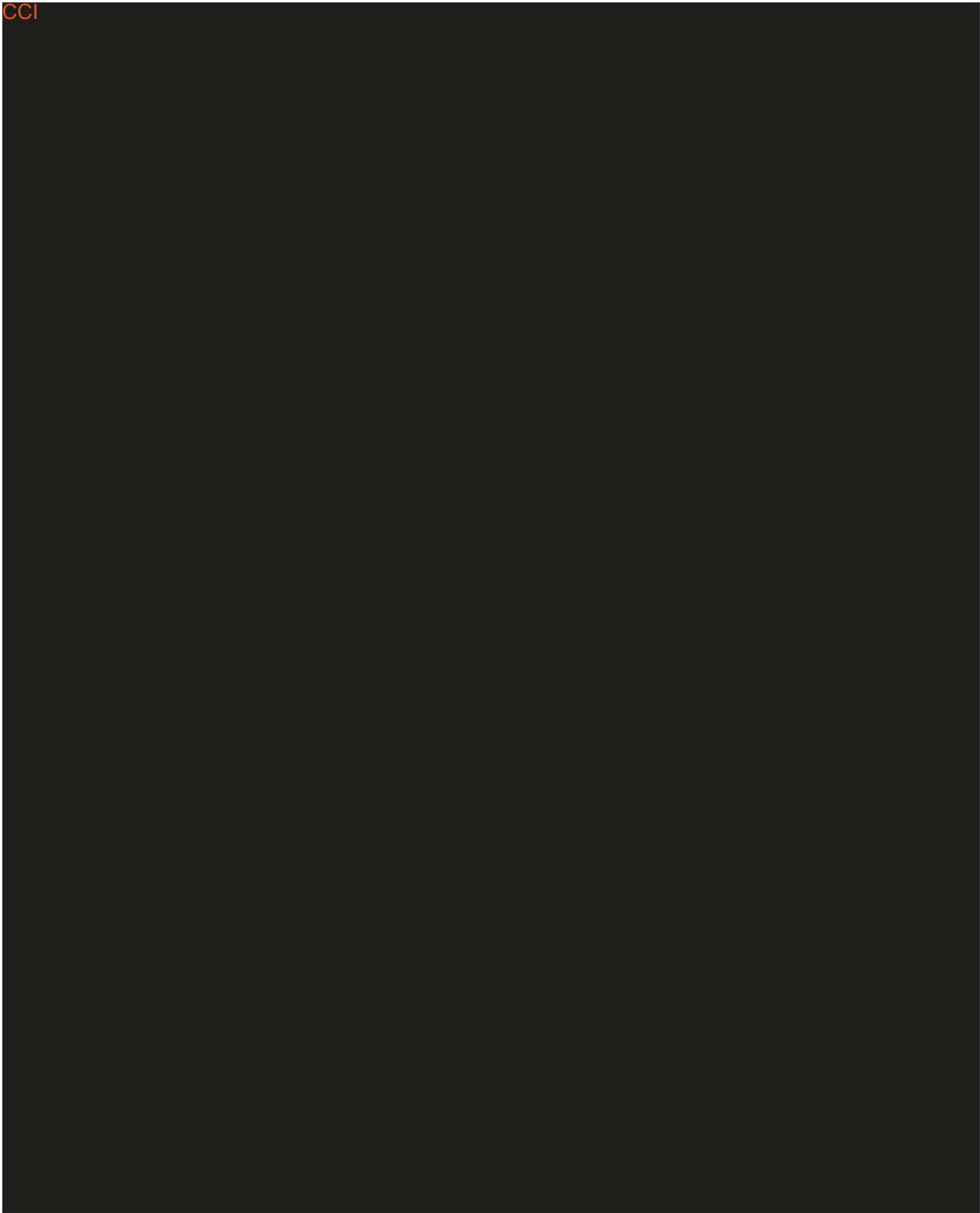
090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

CCI



090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

CCI



3.4. Baseline Variables

Baseline is defined as Day -1 (study derived day and equals to nominal visit Day 0 as noted in the protocol) measurement.

Baseline variables include

- Demographics;
- Medical history;
- Prior/current medication;
- MMSE Total Score;
- Hoehn & Yahr Stage while the subject is ON (if applicable);
- Diary endpoints (OFF time + all other diary measures);
- MDS-UPDRS (Part I, II, III, IV, and Total Score);
- CCI [REDACTED];
- QUIP-RS;
- Primary diagnosis and duration;
- Prescribed L-Dopa dose;
- Background concomitant Parkinson's Disease medications (ATC level classification and a combination of the different classifications);
- Region (stratification factor).

These data will be summarized as part of the baseline characteristics.

The sponsor or designee may verify critical elements of the screening and enrollment process and, in cases where verification is required, will provide written authorization (eg, e-mail) concurring with the investigator assessment that the subject is eligible to return for Day 0 (Randomization) assessments. Eligibility may also be documented via telephone authorization followed by written confirmation. The key elements to be reviewed by the sponsor may include Parkinson's disease diagnosis, MDS-UPDRS scores, concordance testing results, medical history, concomitant medications, and select screening safety assessments including labs, ECGs, and vital signs.

Baseline daily OFF time will be a continuous covariate in the MMRM analysis of the primary analysis (See [Section 5.2.2.1](#)). Geographic region will also be a factor, and defined as North America, Japan, and Europe. The geographic regions are stratification factors for the randomization.

Concomitant and prior medications will be coded using the WHO-drug coding dictionary. In addition, concomitant and prior non-drug treatments/procedures will also be coded using the MedDRA coding dictionary.

3.5. Safety Endpoints

Adverse events, ECGs, blood pressure, pulse rate, C-SSRS, QUIP-RS, PWC-20 and safety laboratory data will be reviewed and summarized on an ongoing basis during the study to evaluate the safety of subjects.

3.5.1. Adverse Events

An adverse event is considered treatment emergent relative to a given treatment if:

- the event occurs for the first time during the effective duration of treatment and was not seen prior to the start of treatment (for example, during the baseline or run-in period), or
- the event was seen prior to the start of treatment but increased in severity during treatment.

The effective duration of treatment is determined by the lag time. Any event occurring within the lag time, whether this occurs during a break in treatment or at the end of treatment, is attributed to the corresponding treatment period. An infinite lag will be used for the study.

A 3-tier approach will be used to summarize AEs. Under this approach, AEs are classified into 1 of 3 tiers. Different analyses will be performed for different tiers (See [Section 6.6.1](#)).

Tier-1 events: These are pre-specified events of clinical importance and are maintained in a list in the product's Safety Review Plan (no Target Medical Events {TMEs} have been defined at this point in time)

Tier-2 events: These are events that are not tier-1 but are "common". A MedDRA PT is defined as a Tier-2 event if there are 4 or more subjects with that PT in any treatment group, given the trial will have less than 400 or fewer subjects per treatment group ("Rule of 4").

Tier-3 events: These are events that are neither Tier-1 nor Tier-2 events. Pfizer standard safety output where all AEs will be included (ie, no new outputs).

Adverse Events will be determined Overall, and separately for Period A and Period B.

3.5.2. Laboratory Data

Laboratory data will be measured at times specified in the STUDY PROCEDURES section of the protocol. Prolactin data collection has the potential to unblind subjects and will be handled blindly by the study team. The Data Blinding Plan will provide details.

Baseline will be the Day 0 measurement.

Determine if there are any laboratory data abnormalities of potential clinical concern as defined in Pfizer Data Standards. This data will be determined Overall, and separately for Period A and Period B.

3.5.3. Vital Signs (Blood Pressure and Pulse Rate)

Blood pressure and pulse rate will be measured at times specified in the STUDY PROCEDURES section of this protocol. Subject temperature must also be assessed at Screening and Day 0 (Randomization) to confirm no symptoms of fever (see Exclusion Criteria #6). Additional collection times, or changes to collection times of blood pressure and pulse rate will be permitted, as necessary, to ensure appropriate collection of safety data.

Orthostatic hypotension is defined as a decrease of ≥ 20 mmHg for systolic blood pressure or ≥ 10 mmHg for diastolic blood pressure 2 minutes after standing from a supine position. Orthostatic hypotension may be symptomatic or asymptomatic.

Baseline will be the Day 0 measurement.

Determine if there is any blood pressure or pulse rate data abnormalities of potential clinical concern as defined in Pfizer Data Standards. This data will be determined Overall, and separately for Period A and Period B.

3.5.4. Electrocardiogram (ECG)

Electrocardiograms (ECGs) should be collected at times specified in the STUDY PROCEDURES section of this protocol.

The average of the triplicate readings collected at each assessment time will be calculated prior to analyzing each ECG parameter. If more than three readings are collected at one triplicate ECG assessment time, average across all readings will be calculated. If any of the three individual ECG tracings has a QTc value ≥ 500 msec, but the mean of the triplicates is not ≥ 500 msec, the data from the subject's individual tracing will be described in a safety section of the study report in order to place the ≥ 500 msec value in appropriate clinical context. However, values from individual tracings within triplicate measurements that are ≥ 500 msec will not be included in the categorical analysis unless the average from the triplicate measurements is also ≥ 500 msec. The mean measurement is reported.

Baseline will be the average of the triplicate ECG measurements collected on Day 0. Change and percent change from baseline will be calculated at each time point for each ECG parameter.

If not supplied, QTcF will be derived using Fridericia's heart rate correction formula:

$$QTcF = QT / (RR)^{1/3} \quad \text{where } RR = 60 \text{ bpm/HR (if not provided)}$$

If QTcB is collected, then it should be listed only.

Determine if there are any ECG data abnormalities of potential clinical concern as defined in Pfizer Data Standards. This data will be determined Overall, and separately for Period A and Period B.

3.5.5. Physical Examination

Physical Examination will be measured at times specified in the STUDY PROCEDURES section of the protocol.

3.5.6. Neurological Examination

Neurological Examination will be measured at times specified in the STUDY PROCEDURES section of the protocol.

3.5.7. Assessment of Suicidal Ideation and Behavior

3.5.7.1. Columbia Suicide Severity Rating Scale (C-SSRS)

The Columbia Suicide Severity Rating Scale (C-SSRS) is an interview based rating scale to systematically assess suicidal ideation and suicidal behavior. Versions are available for "lifetime/Past 12 months" and "since last evaluation". The "lifetime/Past 12 months" evaluation is done at screening, and the "since last evaluation" is done at all other time points.

The C-SSRS should be collected at times specified in the STUDY PROCEDURES section of this protocol by an appropriately trained clinical site staff member. The C-SSRS can also be administered at any time in the study at the discretion of the investigator based on any reasonable concern.

C-SSRS responses will be mapped to the Columbia Classification Algorithm of Suicide Assessment (C-CASA).

Table 3. C-SSRS Mapped to C-CASA - Suicidality Events and Codes

C-CASA Event Code	C-CASA Event	C-SSRS Response
1	Completed suicide	As captured in the safety database
2	Suicide attempt	“Yes” on “Actual Attempt”
3	Preparatory acts towards imminent suicidal behavior	“Yes” on any of the following: <ul style="list-style-type: none"> • ”Aborted attempt”, <u>or</u> • “Interrupted attempt”, <u>or</u> • “Preparatory Acts or Behavior”
4	Suicidal ideation	“Yes” on any of the following: <ul style="list-style-type: none"> • “Wish to be dead”, <u>or</u> • “Non-Specific Active Suicidal Thoughts”, <u>or</u> • “Active Suicidal Ideation with Any Methods (Not Plan) without Intent to Act”, <u>or</u> • “Active Suicidal Ideation with Some Intent to Act, without Specific Plan”, <u>or</u> • “Active Suicidal Ideation with Specific Plan and Intent”
7	Self-injurious behavior, no suicidal intent	“Yes” on “Has subject engaged in Non-suicidal Self-Injurious Behavior?”

The following 3 endpoints are key endpoints for suicidality data analysis and evaluation:

- Suicidal Behavior;
- Suicidal Ideation;
- Suicidal Behavior or Ideation.

Suicidal behavior: A subject is said to have suicidal behavior if the subject has experienced any of the following events (C-CASA event codes 1-3):

- Completed suicide;
- Suicide attempt; or
- Preparatory acts toward imminent suicidal behavior.

Suicidal ideation: Any observed suicidal ideation maps to a single C-CASA category. Depending on the scale used, more granularity of observed ideation (sub-categories of C-CASA category 4) may be displayed. The C-SSRS, for example, includes five ideation questions (that map to C-CASA category 4) with increasing severity.

Subjects with new onset suicidality: A subject will be considered to have a new onset of suicidality if the subject reported no ideation and no behavior at the baseline assessment (note that self-injurious behavior, no suicidal intent [C-CASA code 7] is not considered to be suicidal ideation or behavior) and reported any behavior or ideation post-baseline. Data observed at screening is not considered in the definition of new onset.

Subjects with worsening suicidality relative to baseline: A subject will be considered to have a worsening of suicidality if the subject moved to a lower numbered C-CASA category (observed in categories 1-4) than was reported at baseline. Movement within C-CASA categories 5-9 would not be considered worsening. In addition, worsening will be considered within the suicide ideation C-CASA category 4 if there is an increase in severity identified in the C-SSRS which captures additional granularity on suicide ideation. A subject who reports only ideation at baseline and who reports any behavior post-baseline is considered to have worsened. Data observed at screening is not considered in the definition of worsening.

Table 4 shows examples of new onset suicidality and worsening suicidality for C-SSRS after mapping to C-CASA.

Table 4. C-SSRS Mapped to C-CASA – Examples of Worsening/New Onset

New Onset		Worsening	
Baseline	Any post-baseline (or by visit)	Baseline	Any post-baseline (or by visit)
No ideation and no behavior	C-CASA code=4 (any ideation) <u>or</u> C-CASA code=1, 2, or 3 (any behavior)	Only C-CASA code=4 (Ideation only)	C-CASA code=3, 2, or 1 (any behavior)
		Lowest C-CASA code=4	C-CASA code=3, 2, or 1 (any behavior)
		Lowest C-CASA code=3	C-CASA code=2, or 1
		Lowest C-CASA code=2	C-CASA code=1

With the C-SSRS, worsening may also be observed within suicidal ideation. In this case, the order of suicidal ideations with increasing worsening is as follows: wish to be dead, non-specific active suicidal thoughts, active suicidal ideation with any methods (not plan) without intent to act, active suicidal ideation with some intent to act, without specific plan, and active suicidal ideation with specific plan and intent. All of these values will be considered worsening for reporting.

3.5.7.2. Mental Health Risk Assessment (MHRA)

At each suicidality assessment as per STUDY PROCEDURES, subjects felt to have significant suicidal ideation with actual plan and intent or suicidal behavior, must be evaluated by a clinician/mental health professional (MHP) skilled in the evaluation of suicidality in the subjects by virtue of training or experience (eg, psychiatrist, licensed clinical psychologist) who will determine if it is safe for the subject to participate/continue in the study. Specific criteria that indicate a need for such an assessment are:

- *Suicidal ideation associated with actual intent and/or plan in the past year; (a “YES” answer to C-SSRS questions 4 “some intent to act without specific plan” or 5 “specific plan and intent”).*
- *Previous history of suicide behaviors in the past 10 years (a “YES” answer to any of the suicidal behavior items of the C-SSRS with the behavior occurring in the past 10 years).*
- *In the investigators judgment a risk assessment or exclusion is warranted.*

When there is a positive response to any question on the C-SSRS, the investigator should determine whether an adverse event has occurred.

The PHQ-8 will be performed only at screening. Subjects that meet the Criterion PHQ-8 total score ≥ 15 indicate a need for risk assessment by MHP.

At the baseline visits (Screening, Day 0, or both), a risk assessment will be done by qualified staff at the clinical site to determine whether it is safe for the subject to be enrolled or to continue to participate in the study.

Subjects who respond “YES” to items 4, 5 or to any behavioral question of the C-SSRS at any time after the baseline visit (Screening or Day 0 (Randomization)) will be assessed by clinician/MHP to determine whether it is safe for the subject to continue in the study.

3.5.8. Questionnaire for Impulsive-Compulsive Disorders in Parkinson’s Disease – Rating Scale (QUIP-RS)

The Questionnaire for Impulsive-Compulsive Disorders (ICDs) in Parkinson’s Disease – Rating Scale (QUIP-RS) is a brief, patient reported outcome measure designed to assess the severity of symptoms of ICDs and related behaviors reported to occur in Parkinson’s disease.

The QUIP-RS has 4 primary questions pertaining to commonly reported thoughts, urges/desires, and behaviors associated with impulsive-compulsive disorder, each applied to the 4 impulsive-compulsive disorders (compulsive gambling, buying, eating, and sexual behavior) and 3 related disorders (medication use, punning, and hobbyism). It uses a 5-point Likert scale (score 0–4 for each question) to gauge the frequency of behaviors, and

instructs subjects to answer questions based on behaviors that occurred in the preceding 4 weeks (or any 4-week period in a designated time frame). Each question is anchored with the following five responses: Never (0), Rarely (1), Sometimes (2), Often (3), and Very Often (4). The scoring range for each item (ie, disorder) is 0-16. Scores will be computed for Gambling (0-16), Buying (0-16), Sex (0-16), Eating (0-16) and Total ICD score (0-64). The Total ICD score includes all 4 impulsive-compulsive disorders. Additionally, scores are computed for Hobbyism-punding (combined, 0-32), Compulsive medication use (0-16) and Total QUIP-RS score (0-112). The Total QUIP-RS score includes all 7 disorders. The higher score indicates a greater level of the Impulsive-Compulsive Disorder.

The general principle of allowing up to 20% of missing responses is applied. Therefore, the following subscores are set as missing if any of the pertaining 4 primary questions has a missing response: Each of the four ICDs (compulsive gambling, buying, eating and sexual behavior) and compulsive medication use. For the hobbyism-punding subscore that has 8 items, 1 missing response is allowed. For the Total ICD score that has 16 items, up to 3 missing responses are allowed. For the Total QUIP-RS score that has 28 items, up to 5 missing responses are allowed. This rule allows for the derivation of total scores even when some of the subscores are missing. With the presence of allowable missing responses, any sub or total will be calculated as follows:

$$(\text{Sum of the nonmissing item scores}) \times \frac{\text{total number of items}}{\text{number of nonmissing items}}$$

3.5.9. Physician Withdrawal Checklist (PWC-20)

The PWC-20 is a physician-completed, 20-item reliable and sensitive instrument for the assessment of benzodiazepine discontinuation symptoms. The PWC-20 is collected after the completion of study treatment and also at the first visit of follow-up.

Determine the number of subjects with each symptom present (eg, mild or higher severity) and categorize each subject by severity (eg, mild, moderate, and severe). Only non-missing items are considered in summary presentations, and will establish the denominator.

The total PWC-20 score is the sum of 20 item-scores and ranges between 0 and 60. The higher score indicates more frequent/severe symptoms. If more than 5 of the 20 individual items are missing then the total PWC-20 score will be set to missing.

If more than 5 items are missing, the total PWC-20 score is missing; otherwise, the total PWC-20 score will be imputed as follows: sum of the non-missing items X (total number of items)/ (number of items non-missing).

4. ANALYSIS SETS

Data for all subjects will be assessed to determine if subjects meet the criteria for inclusion in each analysis population prior to unblinding and releasing the database and classifications will be documented per standard operating procedures.

4.1. Full Analysis Set

The dataset used in the efficacy analyses will be the Full Analysis Set (FAS), consisting of all subjects randomized who receive at least 1 post-dose efficacy measurement (Hauser home diary). The FAS is the primary population for efficacy.

All subjects who receive at least 1 post-dose efficacy measurement will be classified according to their randomized treatment assignment. Randomized but took incorrect treatment subjects and randomized but not treated subjects will be reported under their randomized treatment group for all efficacy analyses. Treated but not randomized subjects will be excluded from the efficacy analyses since randomized treatment is missing.

4.2. Per Protocol Analysis Set

The Per Protocol Analysis Set (PPAS) will be a subset of the FAS dataset. This set will exclude subjects based on the following criteria:

- Treatment compliance is outside the 80%-120% compliance range, calculated based on pill count (See [Section 6.5.3](#)).
- Subjects who did not have valid Baseline Hauser home diaries or fail to complete at least Week 5 Hauser diary measurements.
- Subject with any major protocol violations (including but not limited to the violation of entry criteria, use of excluded medication, errors in treatment assignment, loss to follow-up and missing data) that will impact data quality and/or the interpretation of the results.

A full list of protocol deviations will be compiled and reviewed to identify major and minor deviations, and the precise reasons for excluding subjects from the PPAS will be fully defined and documented (Subject Evaluability Determination List) before breaking the blind for the final database release.

4.3. Safety Analysis Set

The safety analysis set will include all patients who received at least one dose of PF-06649751 or placebo. The safety analysis set is the primary population for treatment administration/compliance and safety.

All subjects who receive at least one dose of study medication will be classified according to the actual study treatment received. A randomized but not treated subject will be excluded from the safety analyses. A treated but not randomized subject or a randomized but took incorrect treatment subject will be reported under the treatment actually received.

CCI

5. GENERAL METHODOLOGY AND CONVENTIONS

See the Interim Analysis Plan (IAP) and Internal Review Committee (IRC) Charter for details on any unblinding for the interim analysis.

The blind for the study will be broken and the final analysis of the study data will be conducted once the last remaining subject has completed the study or is withdrawn from the study prior to completion, all data have been entered into the database, all data issues resolved, the Per Protocol Analysis Set (PPAS) has been determined, and the database has been locked.

CCI

5.1. Hypotheses and Decision Rules

5.1.1. Statistical Hypotheses

There are no formal statistical hypotheses in this study generally, and the study decision will be based on the two-part decision rule described in [Section 5.1.2](#) on the primary endpoint daily OFF time in hours at Week 10. To control its overall study Type I error, a preliminary dose-response trend test of the hypothesis of H_0 : no effect for all doses will be performed at the one-sided 0.05 level.

5.1.2. Statistical Decision Rules

The decision criteria for efficacy are given as below, which will be based on a Bayesian interpretation of the results:

- C1: Point estimate of PF-06649751 vs. placebo effect ≥ 1.5 hours, and
- C2: 95% confident of PF-06649751 vs. placebo effect ≥ 0 .

Better is defined as a reduction in the daily OFF time.

No adjustments for the interim analysis or multiple comparisons will be made.

5.2. General Methods

Descriptive summaries for the efficacy, safety, baseline, and CCI will be displayed overall (combining both cohorts) and by cohort/treatment using the following combinations.

- Overall
 - Placebo;
 - PF-06649751 1 mg QD;
 - PF-06649751 3 mg QD;
 - PF-06649751 7 mg QD;
 - PF-06649751 15 mg QD; all 15 mg QD subjects;
 - PF-06649751 15-15 mg QD (safety only); subjects not requiring down titration;
 - PF-06649751 15-7 mg QD (safety only); subjects requiring down titration.
- Cohort 1
 - Placebo
 - PF-06649751 15 mg QD;
 - PF-06649751 15-15 mg QD (safety only);
 - PF-06649751 15-7 mg QD (safety only).
- Cohort 2
 - Placebo;
 - PF-06649751 1 mg QD;
 - PF-06649751 3 mg QD;
 - PF-06649751 7 mg QD;
 - PF-06649751 15 mg QD;
 - PF-06649751 15-15 mg QD (safety only);
 - PF-06649751 15-7 mg QD (safety only).

Note that several visits have been changed as a result of the B7601003 Protocol Amendment 2. Days 14 (Week 2) and 21 (Week 3) were required “in-clinic” visits and now may be handled as phone visits, so those with phone visits will have less data collected. Additionally, Day 91 (Week 13) is now exclusively a phone visit, so only subjects consented under Protocol Amendment 2 will have data available. All available visit data will be reported, and for descriptive statistics by time point each visit will be represented. Additionally, for subjects randomized in Japan, specific “in-clinic” visits (eg, Days 15 and 22) were not previously mentioned in the previous version of the SAP, but will be included in all reporting. Tables and listings will add footnotes as appropriate.

5.2.1. Analyses for Binary Data

Analyses for any binary data output will show the number and percentage of subjects in each response category, and the response rate and its 90% confidence interval (CI) will be calculated. The confidence interval of the response rate will be calculated based on a normal approximation method. Also, the difference of the response rates (with 90% CIs) between each PF-06649751 treatment group and placebo group will be presented.

The analyses for binary data will utilize the logistic regression model (SAS code will be provided in the Programmer's Analysis & Reporting documentation). The response variable of the model will be the binary measurement (CCI [REDACTED]). The model will include the following terms as explanatory variables: geographic region and treatment. The odds ratio and its 90% confidence interval comparing each PF-06649751 dose versus placebo will be presented.

5.2.2. Analyses for Continuous Data

Descriptive statistics n, mean, median, standard deviation, minimum, and maximum will be used to summarize the endpoints. The differences of each PF-06649751 treatment versus the placebo group with 90% confidence intervals from the fitted model will be reported.

5.2.2.1. Mixed Model for Repeated Measures Analysis (Non-Bayesian Approach)

The Mixed Model for Repeated Measures (MMRM) analysis will be based on an analysis of a restricted maximum likelihood (REML)-based mixed model. The response variable will be the change from baseline to each post baseline visit, and the model will include the following fixed effects:

- randomized treatment (with 5 levels: PF-06649751 1 mg QD, 3 mg QD, 7 mg QD, 15 mg QD, and placebo), a categorical factor;
- visit, a categorical factor;
- treatment-by-visit interaction;
- baseline endpoint value, a continuous covariate;
- baseline-by-visit interaction;
- geographic region as appropriate (e.g. North America, Japan, and Europe).

In addition, cohort (Part A, Part B) may also be included as a fixed effect as appropriate.

An unstructured variance-covariance structure will be used to model the within-subject errors. In the unlikely event that the computational algorithm fails to converge, the following structures will be executed in the order specified (essentially in decreasing order of complexity) until convergence is achieved: heterogeneous Toeplitz, heterogeneous first-order

autoregressive, autoregressive, heterogeneous compound symmetry, compound symmetry, and variance components. The first structure yielding convergence will be used as the primary analysis. The Kenward and Roger method⁷ will be used for calculating the denominator degrees of freedom for tests of fixed effects. OM option will be used to derive the least squares means similar to the observed margins. Example SAS code for the MMRM model will be provided in the Programmer's Analysis & Reporting documentation.

The least-squares mean estimates, standard errors, and two-sided 90% CIs for each treatment groups will be reported. The efficacy comparisons based on the treatment difference vs. placebo estimated using least-squares means for each PF-06649751 dose will be conducted, and their point estimates, standard errors, and two-sided 90% CIs will be reported.

The assumptions of normality will be verified graphically using residual plots. For each fitted model, studentized residual plots will be produced. The residual plots will not be included in the clinical study report, but will be sent to the statistician for inspection at the time the analysis is conducted.

5.2.2.2. Dose Response Analysis (Non-Bayesian Approach)

A four parameter E_{\max} model will be used to estimate the dose response relationship. The model is as follows:

$$y = E_0 + \frac{E_{\max} \cdot D^\lambda}{ED_{50}^\lambda + D^\lambda}; \text{ where,}$$

- y is the endpoint change from baseline;
- D is the PF-06649751 dose in mg;
- E_0 is the placebo response;
- E_{\max} is the additive increase over placebo in the response of PF-06649751 at a theoretically infinite dose;
- ED_{50} is the PF-06649751 dose in mg that provides an effect of $E_0 + E_{\max}/2$;
- λ is the Hill parameter and measures the steepness of the dose response curve.

The non-linear dose response model will be fitted in two stages: a MMRM model described in [Section 5.2.2.1](#) will firstly be modelled to the repeated measures data, and the response least squares estimates for the specified visit time as well as the estimated covariance matrix for the response estimates will be obtained. These MMRM estimates will then be fitted to the E_{\max} model, based on the generalized least squares criterion using the inverse of the covariance matrix S as the weighting matrix: $(f(dose, \theta) - resp)' S^{-1} (f(dose, \theta) - resp)$. This fitting is based on the asymptotic normality assumption of the first stage MMRM estimates.

Example R code for the dose response model utilizing DoseFinding package will be provided in the Programmer's Analysis & Reporting documentation.

The efficacy comparisons will be based on the treatment difference vs. placebo estimated from this E_{\max} model fit. Their point estimates and two-sided 90% CIs from the fitted model will be reported.

An acceptable range of a true effect for an minimum effective dose (MED) will be evaluated from the E_{\max} model, and is defined as the lowest PF-06649751 efficacious dose that

- is within the range of tested doses in the study,
- has a lower tested dose that is not efficacious, and
- has a placebo-corrected point estimate between 1.5 and 2 hours.

5.2.2.3. Bayesian Analysis

In Bayesian analysis approach, Bayesian methods will be used to estimate the posterior distribution of the treatment effect. This will be done by combining the data from the study with prior belief about the placebo response.

Informative priors for the placebo response were derived from a Bayesian predictive meta-analysis of a few recent and diverse Parkinson's disease studies. It is possible that whilst the study is ongoing, data information from more sources may become available for incorporation into the informative prior for the placebo effect. The priors specified in this SAP may therefore be amended prior to final study reporting. Any change to the SAP-specified priors will be documented in the CSR.

The informative priors for the placebo effect are defined as

- Prior for the placebo effect for daily OFF time change from baseline $\sim N(-1.0, 0.602)$

A robust mixture prior approach (Bolstad 2007)² will be used for all Bayesian analysis when incorporating informative prior for the placebo effect. The robust mixture informative prior takes into account that should the final data on the endpoint be sufficiently different to the prior perceived knowledge of placebo, then low-informative priors will be used rather than the informative priors. The informative prior is initially weighted with a 90% probability, with a 10% weight to the low-informative prior. The actual weighting of the informative and low-informative priors will be estimated from the data.

A low-informative prior is defined as a distribution with large variance, e.g. $N(0, 10^6)$.

5.2.2.3.1. Bayesian ANOVA Model

In Bayesian ANOVA analysis, treatment will be modelled as a fixed effect, and the least squares estimates and their covariance from MMRM analysis described in Section 5.2.2.1 will be input to the Bayesian analysis. Robust mixture informative prior will be used for the placebo effect, and low-informative prior will be used for all PF-06649751 dose effects. Example R and JAGS code for the Bayesian ANOVA model will be provided in the Programmer's Analysis & Reporting documentation.

The efficacy comparisons will be based on the treatment difference vs. placebo estimated from this Bayesian analysis. Their posterior point estimates and two-sided 90% credible intervals from the fitted model will be reported.

5.2.2.3.2. Bayesian Dose Response Model

In Bayesian dose response analysis, the same four parameter E_{\max} model as described in Section 5.2.2.2 will be used to model the dose response relationship. The least squares estimates and their covariance from MMRM analysis described in Section 5.2.2.1 will be inputted in the Bayesian analysis. Robust mixture informative prior will be used for the placebo effect, and weak empirically-based prior distributions, based on meta-analysis of many dose response studies (Thomas et al 2014), will be used for the other model parameters. These weak priors are defined in this study as:

- Prior for $E_{\max} \sim N(0, 10^6)$;
- Prior for $ED_{50} = \exp(ed50t) * \text{projected } ED_{50}$ (which is 3mg projected for PF-06649751 based on Ph1 studies), where $ed50t \sim t$ distribution $t(0.79, 2.78, 5)$;
- Prior for $\lambda = 6 * lamt$, where $lamt \sim dbeta(3.03, 18.15)$.

Example R and JAGS code for the Bayesian E_{\max} model will be provided in the Programmer's Analysis & Reporting documentation. The efficacy comparisons will be based on the treatment difference vs. placebo estimated from this Bayesian E_{\max} model. Their posterior point estimates and two-sided 90% credible intervals from the fitted model will be reported. MED will be similarly evaluated as described in [Section 5.2.2.2](#).

A Bayes posterior predictive test for the monotonicity of the Emax fit will be conducted by computing the probability that the maximum response from lower doses (all 3 lower doses in this case) compared to the response from the highest dose is no better than predicted by the model (eg, Thomas et al 2014). Example R function code will be provided in the Programmer's Analysis & Reporting documentation. A Bayes predictive probability of < 0.05 will reject the monotonic Emax fit.

5.2.2.4. Dose Response Trend Test

A test of the hypothesis of no effect for all doses will be performed at the one-sided 0.05 level. The MCP-MOD test (Bretz et al 2005)⁴ will be used, with contrasts constructed from the Emax model with candidate ED₅₀ values of 0.5, 1, 3, 7, 15 and 30 mg. Example R function code will be provided in the Programmer's Analysis & Reporting documentation.

5.2.3. Analyses for Categorical Data

Categorical data (eg, CCI and CCI) will show the number and percentage of subjects for each category and include mean/median descriptive statistics similar to the continuous data (See Section 5.2.2). Analysis of the categorical data will use Cochran-Mantel-Haenszel methods, where row-mean-score-difference tests use ridit scores that yield nonparametric analyses. Each PF-06649751 treatment dose will be separately compared to placebo controlling for the geographic region effect.

5.3. Methods to Manage Missing Data

For the analysis of safety endpoints, the sponsor data standard rules for imputation will be applied (eg, partial dates for AEs and concomitant medications will be imputed according to Pfizer standard algorithms).

For the Hauser home diary data, an algorithm is used to handle missing 30-minute intervals (See Section 3.1). A summary of missing diary data will be presented by treatment group.

For scales used in this study, scores will be imputed according to the imputation rules and algorithms for missing component scores that are provided in the data standard documents, or scale documentation. Details are included for each endpoint in Section 3.

Efficacy endpoint analyses in this study will be primarily based on observed cases (OC) and missing data will not be imputed, except for CCI, in which case last observation carried forward (LOCF) will be utilized.

Efficacy and safety analyses will be based on nominal visits as recorded on the CRFs. The only exception will be the early termination (ET) visit if a subject terminates before completion of Week 15. If the subject terminates before Week 15, the data collected at the early termination visit will be assigned to the closest planned visit (target day) for that efficacy or safety assessment. If the early termination visit is equidistant to 2 visits, it will be assigned to the later visit. If the early termination visit is associated with a visit that already exists, the non ET visit will be the representative visit. However, for purposes of LOCF, if more than one value is in a visit window the latter value will be used to carry forward. For all safety categorical determinations, early termination and unplanned visits are considered even if they are not assigned to a visit.

If a subject is missing Baseline (eg, Day 0), the last measurement, taken prior to first exposure to the investigational product, will be used.

Concentrations below the limit of quantification

In all data presentations (except listings), concentrations below the limit of quantification (BLQ) will be set to zero. (In listings BLQ values will be reported as “<LLQ”, where LLQ will be replaced with the value for the lower limit of quantification).

Deviations, missing concentrations and anomalous values

In summary tables and plots of median profiles, statistics will be calculated having set concentrations to missing if 1 of the following cases is true:

1. A concentration has been collected as ND (ie, not done) or NS (ie, no sample);
2. A deviation in sampling time is of sufficient concern or a concentration has been flagged anomalous by the pharmacokineticist.

Note that summary statistics will not be presented at a particular time point if more than 50% of the data are missing.

6. ANALYSES AND SUMMARIES

Both the FAS and PPAS populations will be used in the analyses of the primary efficacy endpoint, with the FAS being primary. For all other efficacy endpoints, only the FAS analysis set will be utilized.

For all efficacy analyses and summaries (Sections 6.1, 6.2, 6.3 {except CCI [REDACTED]} and 6.4): All data collected for subjects consented under Protocol Amendment 2 prior to entering Period B will be included for modelling and summaries. Otherwise, only data collected through Week 10 (Period A) will be included for subjects who do not meet the above condition, and their data collected after Week 10 will be in listings only.

6.1. Primary Endpoint(s)

6.1.1. Change from Baseline in Daily OFF time at Week 10 (end of Period A)

6.1.1.1. Primary Analysis

Endpoint: Daily OFF time

- Analysis time points: Week 10 (End of Period A)
- Analysis population (method of imputation for missing data): FAS (See Section 3.1)
- Analysis methodology

Change from baseline will primarily be analyzed using the Bayesian Dose Response Analysis (See Section 5.2.2.3.2). If the Bayes posterior predictive test rejects the monotonic Emax fit (prob < 0.05), then the primary analysis will be based on the MMRM analysis (See Section 5.2.2.1). An overall dose-response trend test (See Section 5.2.2.4) will also be performed to control family-wise error rate.

- Supporting objective and Decision rule: Primary Objective (See [Section 2.1.1](#)), Secondary Objective (See [Section 2.1.2](#)), and Decision rules C1, C2 (See [Section 5.1.2](#))

Reporting results:

- Raw data: The sample size, mean, standard deviation, median, minimum and maximum at baseline and post-baseline visits will be presented for each treatment arm.
- Change from baseline: The sample size, mean, standard deviation, median, minimum and maximum will be presented for each treatment arm. The Bayesian point estimates, 90% credible interval, difference from placebo for each pair of treatment groups and the corresponding 90% credible interval will be presented. The MED estimate if applicable will be presented.

Figures

- Vertical bar chart of Change from Baseline Bayesian point estimates and 90% credible interval at Week 10.
- Bayesian Emax model graph with the MMRM estimates.
- Change from baseline time profile per MMRM for each treatment group.

6.1.1.2. Sensitivity/Robustness Analyses

To support the interpretation of the primary analysis the following analyses will be performed:

Endpoint: Daily OFF time

- Analysis time points: Week 10 (End of Period A);
- Analysis populations (method of imputation for missing data): FAS, PPAS, FAS-Cohort 1 only, FAS-Cohort 2 only (See [Section 3.1](#));
- Analysis methodology: Change from baseline will be analyzed using the MMRM analysis (See [Section 5.2.2.1](#); FAS, FAS-Cohort 1 only), Bayesian ANOVA analysis (See [Section 5.2.2.3.1](#); FAS, FAS-Cohort 1 only), Dose Response analysis (See [Section 5.2.2.2](#); FAS, FAS-Cohort 2 only), and Bayesian Dose Response analysis (See [Section 5.2.2.3.2](#); PPAS, FAS-Cohort 2 only).

Reporting results:

- Raw data: The missing data pattern will be shown by treatment group and overall, for the Hauser Diary up to Week 10.
- Change from baseline: The point estimate, 90% confidence or credible interval for the point estimate, difference from placebo for each pair of treatment groups and the corresponding 90% confidence or credible interval will be presented. The MED estimate if applicable will be presented.

Figures

- Empirical cumulative distribution function showing % of subjects with change from baseline \leq cutoff value at Week 10 by treatment group; (any drop outs are assigned to the left side of x axis).

6.2. Secondary Endpoint(s)

6.2.1. Hauser Home Diary

Endpoints: Daily OFF time, Daily ON time with troublesome dyskinesia and Daily ON time without troublesome dyskinesia

- Analysis time points: All Visits;
- Analysis population (method of imputation for missing data): FAS (See [Section 3.1](#));
- Analysis methodology: Change from baseline will be analyzed using the MMRM Analysis (See [Section 5.2.2.1](#)) and Dose Response Analysis (Daily ON time without troublesome dyskinesia at Week 10 only, See [Section 5.2.2.2](#));
- Supporting objective and Decision rule: Primary Objective (See [Section 2.1.1](#)), Secondary Objective (See [Section 2.1.2](#))

Reporting results:

- Raw data: The sample size, mean, standard deviation, median, minimum and maximum at baseline and post-baseline visits will be presented for each treatment arm.
- Change from baseline: The sample size, mean, standard deviation, median, minimum and maximum will be presented for each treatment arm. The point estimate, 90% confidence interval for the point estimate, difference from the placebo for each pair of treatment groups and the corresponding 90% confidence interval will be presented.

Figures (Daily ON time with troublesome dyskinesia, without troublesome dyskinesia)

- Line graphs of LS means and 90% confidence interval at all post-baseline visits.
- Vertical bar chart of LS means and 90% confidence intervals at Week 10.
- Empirical cumulative distribution function showing % of subjects with change from baseline \leq cutoff value at Week 10 by treatment group; (any drop outs are assigned to the left side of x axis)
- Dose response Emax model graph with the MMRM estimates at Week 10.

6.2.2. MDS-UPDRS

Endpoints: MDS-UPDRS Parts I, II, III, IV, and total score

- Analysis time points: All Visits
- Analysis population (method of imputation for missing data): FAS (See [Sections 3.2.1 and 3.2.2](#)).
- Analysis methodology: Change from baseline will be analyzed using the MMRM Analysis (See [Section 5.2.2.1](#)) and Dose Response Analysis (MDS-UPDRS Part III @ Week 10 only, See [Section 5.2.2.2](#)).

- Supporting objective and Decision rule: Secondary Objective (See [Section 2.1.2](#)).

Reporting results:

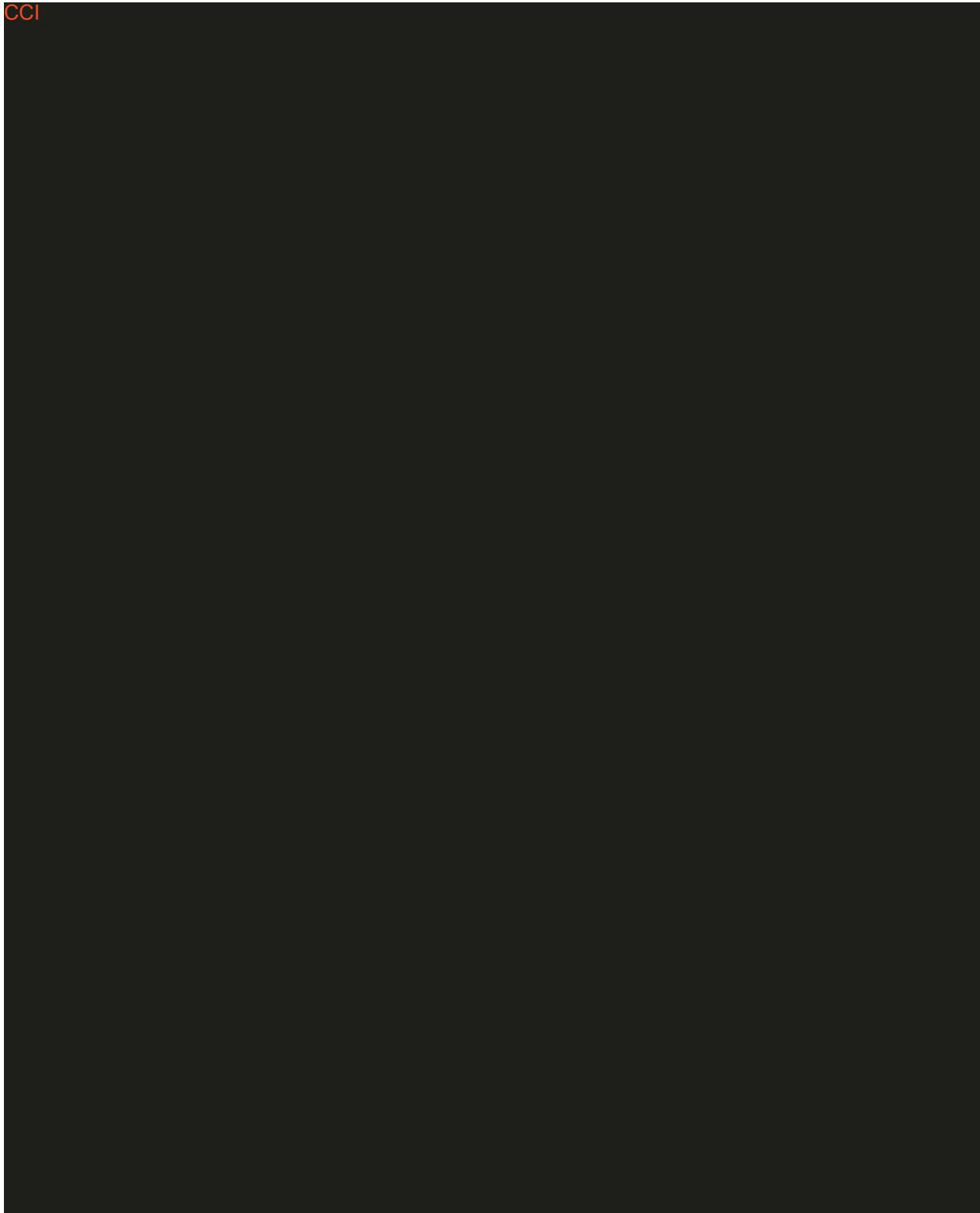
- Raw data: The sample size, mean, standard deviation, median, minimum and maximum at baseline and post-baseline visits will be presented for each treatment arm.
- Change from baseline: The sample size, mean, standard deviation, median, minimum and maximum will be presented for each treatment arm. The point estimate, 90% confidence interval for the point estimate, difference from the placebo for each pair of treatment groups and the corresponding 90% confidence interval will be presented.

Figures

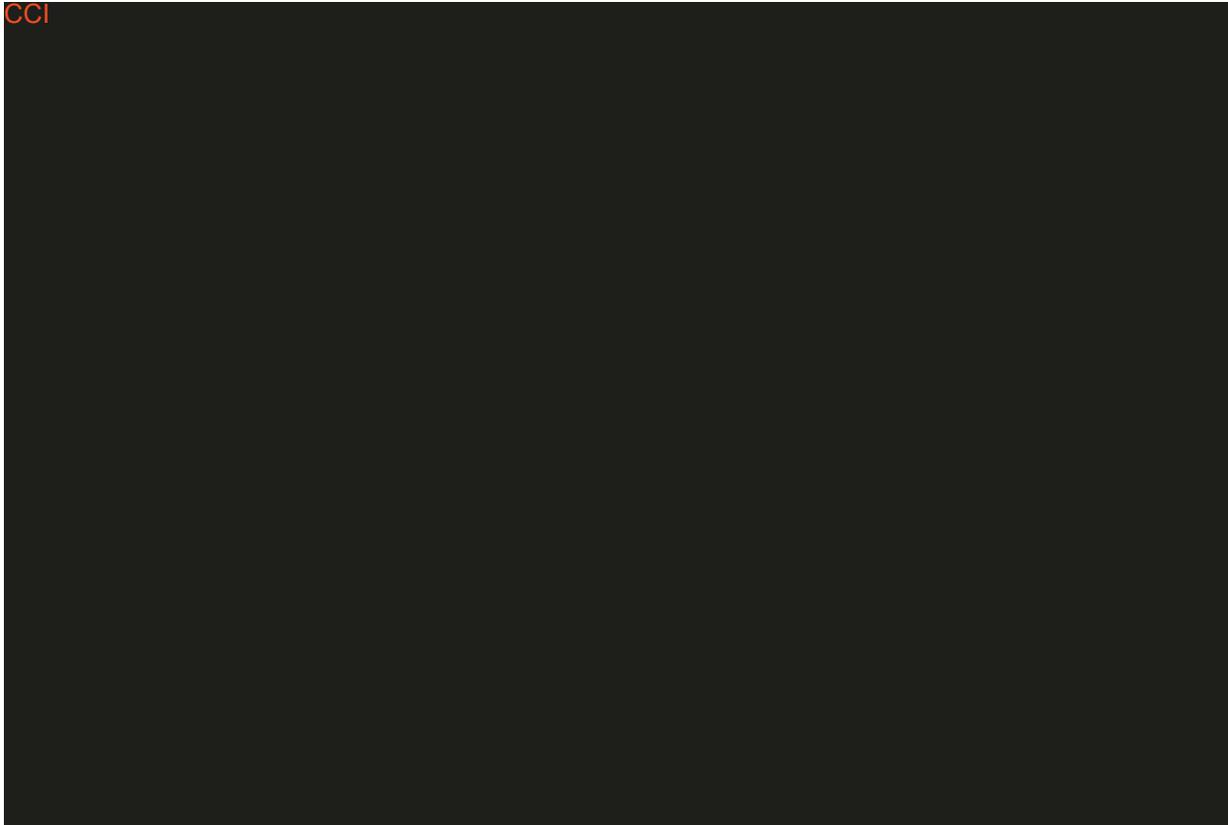
- Vertical bar chart of LS means and 90% confidence intervals at Week 10.
- Change from baseline time profile per MMRM for each treatment group.
- Empirical cumulative distribution function showing % of subjects with change from baseline \leq cutoff value at Week 10 by treatment group; (any drop outs are assigned to the left side of x axis).
- Dose response Emax model graph with the MMRM estimates (MDS-UPDRS Part III at Week 10 only).



CCI



CCI



6.4. Subset Analyses

The primary endpoint will also be analyzed respectively for subset of subjects

- in different geographic regions;
- that have or do not have background PD medication changes during Period A; and
- subjects that have <1000 mg/day or \geq 1000 mg/day of L-Dopa intake during baseline.

Only subgroup that has \geq 10% of total subjects will be analyzed. If conducted, the MMRM analysis (See [Section 5.2.2.1](#)) will be performed and similar reporting results as described in [Section 6.1.1](#) will be presented for each of the subset analyses.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Study Conduct and Subject Disposition

Data will be reported in accordance with Pfizer Data Standards. Disposition will be further categorized by the treatment phase Titration, Dose Adjustment, Period A, Period B (may be separated by preA2 and A2), and Follow-up.

Overall subject disposition will include data through Double-Blind treatment period (study treatment period) and separately through Follow-up (overall study).

The number of subjects by region and country will be displayed.

6.5.2. Baseline Summaries

A breakdown of demographic data will be provided for age, race, weight, body mass index, height, and temperature. Each parameter will be summarized with tables and listings presented in accordance with the Pfizer Data Standards. Use geriatric age categories for the demographic summaries (<65, 65-74, 75-84, and ≥85). Physical measurements at Baseline will include weight, height, BMI, and temperature.

Also, medical history and primary diagnosis will be tabulated and listed in accordance with the Pfizer Data Standards.

Females of childbearing potential are excluded from the study. All female subjects of non-childbearing potential will have a FSH test, if amenorrheic for at least 12 consecutive months. Results will be obtained prior to PF-06649751 or placebo dosing. Information will be captured in the listings only.

The following baseline characteristics will be summarized:

- Hauser home diary endpoints (OFF time + all other diary measures) (continuous data);
- CCI [REDACTED]
- QUIP-RS;
- MMSE Total Score (continuous data);
- Hoehn & Yahr Stage while the subject is ON (categorical data);
- MDS-UPDRS Part I, II, III, IV, and Total Score at Baseline (continuous data);
- Prescribed L-Dopa dose (continuous data);
- Number of Background concomitant Parkinson's Disease medications @ Randomization (categorical);
- Region (categorical).

6.5.3. Study Treatment Exposure

Duration of exposure will be presented in tables and listings in accordance with the Pfizer Data Standards.

Compliance will be calculated and summarized across the whole double-blind treatment phase, and by visit. Compliance will be based on the dosing record, and calculated as:

$$\% \text{Compliance} = 100 \times (\text{Actual \# of Capsules Taken}) / (\text{Total \# Capsules Prescribed})$$

Subjects who fail to return blister information will be considered missing for that visit.

6.5.4. Concomitant Medications and Non-Drug Treatments

All prior and concomitant medication(s) as well as non-drug treatment(s) will be provided in tables and listings in accordance with the Pfizer Data Standards. Concomitant Parkinson's Disease medication(s) use will also be summarized by treatment group.

Frequency and type of medication will be summarized categorically by treatment group. Descriptive statistics of doses of background antipsychotic medication will also be summarized.

The number of rescue medication doses used for subjects will be presented by treatment group overall and by visit (summarized as continuous data).

The number and percentage of subjects with a change in their background PD medications (excluding L-Dopa) for Period A and Period B (summarized as categorical data).

6.6. Safety Summaries and Analyses

The safety analysis set is the primary population for the safety summaries and analyses.

Safety will be summarized Overall, and separately for Period A and Period B for data that is not presented by time point (eg, AEs, concomitant medications, Lab/VIS/ECG categorical determinations).

6.6.1. Adverse Events

Adverse events will be listed and summarized within treatment group in accordance with the Pfizer Data Standards.

The details of Tier-1, Tier-2 and Tier-3 AEs are described in [Section 3.5.1](#).

Adverse events within Tier-1 and -2 will be summarized using Risk Differences between each PF-06649751 group and placebo, together with 95% CI. A graphical presentation of the percentage of subjects with each AE and Risk Difference (with 95% CI) ordered by decreasing risk difference will be shown for the Tier-1 and -2 AEs separately. Significance tests will be performed for the Tier-1 adverse events. There will be no multiplicity adjustment for these significance tests.

For Tier-1 and Tier-2 adverse event outputs, the following footnotes will be used: “P-values and confidence intervals are not adjusted for multiplicity and should be used for screening purpose only. The 95% Confidence intervals are provided to help gauge the precision of the estimates for Risk Difference. Risk Difference is computed as PF-06649751 1 mg versus Placebo, PF-06649751 3 mg versus Placebo, PF-06649751 7 mg versus Placebo, and PF-06649751 15 mg versus Placebo.”

Tier-1 events for risk difference and relative risk will use unconditional exact methods (an approach proposed by Chan and Zhang, 1999⁶) based on standardized statistics. For Tier-2 events, an asymptotic approach will be performed.

The Tier-3 adverse events will be described as part of the overall AE summary.

It should be recognized that most studies are not designed to reliably demonstrate a causal relationship between the use of a pharmaceutical product and an adverse event or a group of adverse events. Except for select events in unique situations, studies do not employ formal adjudication procedures for the purpose of event classification. As such, safety analysis is generally considered as an exploratory analysis and its purpose is to generate hypotheses for further investigation. The 3-tier approach facilitates this exploratory analysis.

6.6.2. Laboratory Data

Laboratory data will be listed and summarized within treatment group in accordance with the Pfizer Data Standards.

Incidence of laboratory test abnormalities (including without regard to baseline abnormality) will be summarized within each treatment group.

6.6.3. Vital Signs (Blood Pressure and Pulse Rate)

For each planned time point, baseline values and change from baseline values within each treatment will be summarized with descriptive statistics (using Pfizer Data Standards). A plot of individual blood pressure versus plasma concentration will be generated.

Maximum decrease and increase values and changes from baseline for vital signs (for supine and standing) will also be summarized descriptively within treatment group using categories as defined in the Pfizer Data Standards. Numbers and percentages of subjects meeting the categorical criteria will be provided. All planned and unplanned post-dose time points will be counted in these categorical summaries. All values meeting the criteria of potential clinical concern will be listed.

Vital signs collected at additional positions will be listed only.

The following non-standard safety tables will be included:

- A summary of postural change from supine to standing systolic and diastolic blood pressures.
- Incidence of subjects with orthostatic hypotension (defined in [Section 3.5.3](#) above), for each visit, last visit and any post-baseline incidence or orthostatic hypotension or minimum absolute change in postural blood pressure.

6.6.4. Electrocardiogram

For each planned time point, baseline values, absolute values and change from baseline values within each treatment will be summarized with descriptive statistics for each ECG parameter (using Pfizer Data Standards).

A plot of QTcF versus plasma concentration will be generated at nominal time points.

Maximum decrease and increase values and changes from baseline for the ECG parameters QT interval, heart rate, QTc interval, PR interval and QRS interval will be summarized by treatment and time post dose using Pfizer Data Standards.

The number (%) of subjects with maximum post dose QTc values and maximum increases from baseline in the following categories will be tabulated by treatment:

Table 5. Safety QTcF¹

	Borderline (msec)	Prolonged (msec)
Absolute Value	≥450 - <480	≥480
Absolute Change	30-<60	≥60

In addition, the number of subjects with corrected and uncorrected QT values ≥500 msec will be summarized.

ECG endpoints and changes from baseline (QTcF, PR, QRS) will also be summarized descriptively by treatment using categories as defined in Pfizer Data Standards (for QTc, these correspond to ICH E14). Numbers and percentages of subjects meeting the categorical criteria will be provided. All planned and unplanned post dose time points will be counted in these categorical summaries. All values meeting the criteria of potential clinical concern will be listed.

6.6.5. Physical Examination

Physical examination data will be listed and summarized within treatment group in accordance with the Pfizer Data Standards.

6.6.6. Neurological Examination

Neurological examination data will be presented in the listings.

6.6.7. Assessment of Suicidal Ideation and Behavior

6.6.7.1. Columbia Suicide Severity Rating Scale (C-SSRS)

In general, the denominator used in the percentages will be the number of subjects assessed for suicidality or worsening, the denominator would include the subset of subjects who had any level of suicidality reported at baseline. For new onset, the denominator would include the subset of subjects with no suicidality reported at baseline.

A subject listing of C-CASA categories as well as the underlying C-SSRS scale data will be presented.

In addition, a summary table with the number and percent of subjects within each C-CASA category by treatment group at screening, baseline, and at any time post-baseline without regard to baseline will be reported.

6.6.7.2. Mental Health Risk Assessment (MHRA)

A summary table by visit will indicate the number and percent of subjects that met the criteria for MHRA evaluation, whether the MHRA module was performed, and if the subject was either eligible to enter the study (Screening) or eligible to continue in the study (All visits after Screening). The category of subjects that met the criteria for MHRA evaluation will include additional details on the criteria met:

- Screening:
 - PHQ-8 total score ≥ 15 ,
 - Subject answered "YES" on C-SSRS item 4, 5, or any suicidal behavior question on the C-SSRS,
 - Active suicidal ideation (with method and/or plan and/or intent) in the past year,
 - Previous history of suicidal behavior in the past 10 years, and
 - Principal Investigator's Discretion.
- All Visits after Screening:
 - Subject answered "YES" on C-SSRS item 4, 5, or on any suicidal behavior question on the C-SSRS,
 - Principal Investigator's Discretion.

6.6.8. Questionnaire for Impulsive-Compulsive Disorders in Parkinson's Disease – Rating Scale (QUIP-RS)

For each planned time point and item (ie, disorder), baseline values, absolute values and change from baseline values within each treatment will be summarized with descriptive statistics for the 8 different scoring types. Additionally, the number and percentage of subjects with each of the 8 different scoring types will be presented.

6.6.9. Physician Withdrawal Checklist (PWC-20)

Summaries of the count and percentage of patients experiencing each symptom and severity listed in the PWC-20 will be provided by treatment group. Follow the PDS used for reporting incidence and severity of Adverse Events.

The Total PWC-20 score will be presented by treatment group using continuous summary statistics for the raw data.

7. INTERIM ANALYSES

7.1. Introduction

Up to two interim analyses may be performed when approximately 27% and 80% subjects are randomized and have the opportunity to complete through Period A of the study. The first interim analysis will be conducted when Cohort 1 (approximately 54 subjects in a 2:1 ratio of 15 mg QD and placebo groups respectively) is completed through Period A and efficacy data have been collected, databased and cleaned. A second interim analysis may be undertaken when approximately 104 subjects complete through Period A of Cohort 2 (1:2:2:2:1 in 15 mg QD, 7 mg QD, 3 mg QD, 1 mg QD, and placebo respectively) and efficacy data have been collected, databased and cleaned. The purpose of these interim analyses will be to assess study futility and to aid future development planning. If an interim analysis is performed, the study will continue whilst the analysis is being conducted.

Unblinded safety data of the study will be periodically reviewed by an external, independent data safety monitoring committee (EDMC). The EDMC will be responsible for ongoing monitoring of the safety of subjects in the study according to the charter. The recommendations made by the EDMC to alter the conduct of the study will be forwarded to Pfizer for final decision. Pfizer will forward such decisions, which may include summaries of aggregate analyses of endpoint events and of safety data that are not endpoints, to regulatory authorities, as appropriate.

7.2. Interim Analyses and Summaries

Before any interim analysis is initiated, the details of the objectives, decision criteria, unblinding, dissemination plan and method of maintaining the study blinded as per Pfizer's standard operating procedures (SOPs) will be documented and approved in an internal review committee (IRC) charter. In addition, the analysis details will be documented and approved in the interim analysis SAP.

8. REFERENCES

1. ICH E14 - The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. CHMP/ICH/2/04.
2. W. Bolstad (2007) Introduction to Bayesian Statistics: Second Edition London: John Wiley & Sons.
3. N. Thomas, K. Sweeney, and V. Somayaji. Meta-analysis of clinical dose-response in a large drug development portfolio. *Statistics in Biopharmaceutical Research*, 6(4): 302-317, 2014.
4. F. Bretz, J. Pinheiro, and M. Branson (2005). Combining multiple comparisons and modeling techniques in dose-response studies. *Biometrics* 61, 738-748.
5. Goetz, et al. Handling Missing Values in the MDS-UPDRS. *Movement Disorders* (2015).
6. Chan ISF, Zhang Z. (1999). Test-based exact confidence intervals for the difference of two binomial proportions. *Biometrics*, 55:1201–1209.
7. Kenward, M. G., and Roger, J. H. (1997). “Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood.” *Biometrics* 53:983–997.

9. APPENDICES**Appendix 1. SUMMARY OF EFFICACY ANALYSES**

Endpoint	Analysis Set	Statistical Method	Missing Data	Interpretation
Daily OFF time, Week 10	FAS	Bayesian Dose Response Analysis. If the Bayes posterior predictive test rejects the monotonic Emax fit (prob < 0.05), then the primary analysis will be based on the MMRM analysis.	OC	Primary Analysis
Daily OFF time, Week 10	FAS	MMRM	OC	Sensitivity Analysis
Daily OFF time, Week 10	FAS	Bayesian ANOVA	OC	Sensitivity Analysis
Daily OFF time, Week 10	FAS	Dose Response	OC	Sensitivity Analysis
Daily OFF time, Week 10	PPAS	Bayesian Dose Response	OC	Sensitivity Analysis
Daily OFF time, Week 10	FAS-Cohort 1	MMRM	OC	Sensitivity Analysis
Daily OFF time, Week 10	FAS-Cohort 1	Bayesian ANOVA	OC	Sensitivity Analysis
Daily OFF time, Week 10	FAS-Cohort 2	Dose Response	OC	Sensitivity Analysis
Daily OFF time, Week 10	FAS-Cohort 2	Bayesian Dose Response	OC	Sensitivity Analysis
Daily OFF time, All Visits	FAS	MMRM	OC	Secondary Analysis
Daily ON time with troublesome	FAS	MMRM	OC	Secondary Analysis

dyskinesia, All Visits				
Daily ON time with troublesome dyskinesia, Week 10	FAS	Dose Response	OC	Secondary Analysis
Daily ON time without troublesome dyskinesia, All Visits	FAS	MMRM	OC	Secondary Analysis
Daily ON time without troublesome dyskinesia, Week 10	FAS	Dose Response	OC	Secondary Analysis
MDS-UPDRS Parts III, All Visits	FAS	MMRM	OC	Secondary Analysis
MDS-UPDRS Parts III, Week 10	FAS	Dose Response	OC	Secondary Analysis
MDS-UPDRS Parts I, II, IV, and total score, All Visits	FAS	MMRM	OC	Secondary Analysis
CCI				
Daily OFF time, Week 10	FAS-North America (Optional*)	MMRM	OC	Subset Analysis

090177e18f6e873a\Approved\Approved On: 26-Sep-2018 05:14 (GMT)

Daily OFF time, Week 10	FAS-EU (Optional*)	MMRM	OC	Subset Analysis
Daily OFF time, Week 10	FAS-Japan (Optional*)	MMRM	OC	Subset Analysis
Daily OFF time, Week 10	FAS-no background PD medication change (Optional*)	MMRM	OC	Subset Analysis
Daily OFF time, Week 10	FAS-background PD medication change (Optional*)	MMRM	OC	Subset Analysis
Daily OFF time, Week 10	FAS-<1000 mg /day and ≥1000 mg/day L-dopa intake (Optional*)	MMRM	OC	Subset Analysis

* Only subgroups that have ≥ 10% of total subjects will be analyzed.

CCI

