STATISTICAL ANALYSIS PLAN

Study: PD0053

Product: UCB0599

A DOUBLE-BLIND, PLACEBO-CONTROLLED, RANDOMIZED 18-MONTH PHASE 2A STUDY TO EVALUATE THE EFFICACY, SAFETY, TOLERABILITY, AND PHARMACOKINETICS OF ORAL UCB0599 IN STUDY PARTICIPANTS WITH EARLY PARKINSON'S DISEASE

PHASE 2A

Sponsor Name: UCB Biopharma SRL

Legal Registered Address: Allée de la Recherche 60

1070 Brussels

BELGIUM

Regulatory Agency Identifier Number(s):

Eudra CT Number:	2020-003265-19
IND Number:	141003

TABLE OF CONTENTS

VERSION HISTORY 5
LIST OF ABBREVIATIONS 29
1 INTRODUCTION 33
1.1 Objectives and estimands/endpoints 33
1.2 Intercurrent events 40
1.2.1 Symptomatic treatment initiation 45
1.2.2 Treatment discontinuation and treatment-related ICE-like protocol deviations 46
1.2.3 Other ICE-like protocol deviations 46
1.2.4 Death or serious injury 46
1.2.5 Confirmed or suspected COVID-19 47
1.2.6 Study termination and loss to follow-up 47
1.3 Study design 48
2 STATISTICAL HYPOTHESES 49
3 SAMPLE SIZE DETERMINATION 49
3.1 Primary efficacy estimand 49
3.2 Incorporating data loss 51
3.3 Power for key secondary efficacy estimand (MDS-UPDRS Part III) 53
3.4 Power for secondary efficacy estimand (DaT-SPECT mean striatum SBR) 53
4 POPULATIONS FOR ANALYSIS 54
5 STATISTICAL ANALYSES 55
5.1 General considerations 55
5.1.1 Date imputation 56
5.1.2 General study level definitions 58
5.1.2.1 Analysis time points 58
5.1.2.2 Protocol deviations 60
5.1.2.3 Treatment assignment and treatment groups 61
5.1.2.4 Center pooling strategy 62
5.1.2.5 Coding dictionaries 62
5.1.2.6 Multicenter studies 62
5.2 Participant dispositions 63
5.3 Efficacy estimands and endpoints 64
5.3.1 Primary estimand analysis: MDS-UPDRS Part I-III sum score slope of progression over 12 months, in the absence of ST initiation 66
5.3.1.1 Definition of endpoint 67
5.3.1.2 Main analytical approach 67
5.3.1.3 Sensitivity analyses 68
5.3.1.4 Supplementary approaches 68

5.3.2 Secondary estimands/endpoint analyses 69
5.3.2.1 Secondary estimands: clinical symptoms (MDS-UPDRS) 69
5.3.2.2 Secondary estimands: clinical symptoms (MoCA) 79
5.3.2.3 Secondary estimands: neurodegeneration 80
5.3.2.4 Secondary estimands: ST Initiation/Intake 82
5.3.3 Exploratory endpoints analysis 84
5.3.3.1 Exploratory endpoints: clinical symptoms 84
5.3.3.2 Exploratory neurodegeneration endpoints 88
5.3.3.3 Exploratory ST intake endpoints 88
5.4 Safety analyses 89
5.4.1 Extent of exposure 89
5.4.2 Adverse events 90
5.4.2.1 Adverse events of special interest (AESI) 92
5.4.3 Additional safety assessments 92
5.4.3.1 Clinical safety laboratory assessments 92
5.4.3.2 Vital signs 95
5.4.3.3 Electrocardiograms 96
5.4.3.4 Physical and neurological examination 97
5.4.3.5 Suicidal risk monitoring 98
5.4.3.6 MRI assessments 98
5.4.3.7 Wearable sensor data 98
5.5 Other analyses 98
5.5.1 Pharmacokinetics 98
5.5.2 Pharmacodynamics 99
5.5.3 Biomarkers 100
5.6 Subgroup analyses 100
5.7 Interim Analyses 100
5.8 Planned analysis at 12 months 100
5.9 Data Monitoring Committee (DMC) or other review board 101
6 SUPPORTING DOCUMENTATION 102
6.1 Appendix 1: Non-key analysis specifications 102
6.1.1 Baseline characteristics and demographics 102
6.1.1.1 Demographics 102
6.1.1.2 Baseline disease characteristics 102
6.1.1.3 Other baseline characteristics 103
6.1.2 Protocol deviations 103
6.1.3 Medical history 103
6.1.4 Prior/concomitant medications 103

6.1.5 Data derivation rules 104
6.1.6 AEs of special interest 104
6.1.7 Potentially clinically significant criteria for safety endpoints 105
6.1.8 Compliance 105
6.2 Appendix 2: Changes to protocol-planned analyses 106
6.3 Appendix 3: Data loss to ST initiation modelling 106
6.4 Appendix 4: Producing summaries by treatment group and overall 109
6.5 Appendix 5: MDS-UPDRS Part I-III sum score, MDS-UPDRS Part III subscale and ePD subscore: imputation, ST de-mediation and modelling 110
6.5.1 Overview of Estimands and General Approach 110
6.5.2 Data Pre-processing 112
6.5.3 Bootstrapping 113
6.5.4 Handling of Missing Data 113
6.5.5 ST de-mediation via g-estimation 115
6.5.6 End-of-trial Analysis Models 117
6.5.7 Results Summaries 118
6.6 Appendix 6: Key inclusion/exclusion criteria for the FAS definition 120
7 REFERENCES 121

LIST OF TABLES

Table 1–1: Study objectives 36
Table 1–2: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for MDS-UPDRS estimands 40
Table 1–3: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for other estimands 43
Table 3–1: Parameter estimates based on PPMI data 51
Table 5–1: Calculation rules for duration of adverse events 57
Table 5–2: Treatment group descriptions 61
Table 5–3: Clinical laboratory assessments 93
Table 5–4: TEMA/PCS criteria for vital signs 96
Table 6–1: Treatment groups to be presented in summary tables 109
Table 6–2: Items to be included in the 15-item MDS-UPDRS Part III ePD subscore 112

LIST OF FIGURES

Figure 6–1: Observed data loss in PPMI dataset 107
Figure 6–2: Simulated data loss in PD0053 108

VERSION HISTORY

SAP Version	Approval Date
Original SAP	19 May 2022
Amendment 1	25 Jan 2023
Amendment 2	10 Oct 2023
Amendment 3	18 Jun 2024
Amendment 4	21 Oct 2024

Amendment 4 (21 Oct 2024)

Overall Rationale for the Amendment

To add in the statistical details for the ST de-mediation analyses.

Section #	Description of Change	Brief Rationale
Section 1.1	Table 1-1 footnotes updated to remove the statement that ST washout compliance is required for 'regardless of ST initiation' analyses. Table 1-1 footnote updated to remove the note that MDS-UPDRS Part I analyses regardless of ST are considered supplementary in the SAP.	Washout compliance requirement removed as there is no clear impact on MDS-UPDRS Part III score with/without washout. We therefore choose to use the data collected for participants who do not comply. SAP aligned with protocol, where these are considered secondary estimands.
Section 1.2	Footnote g added for treatment-related study termination.	To clarify that ST de-mediation approaches may implement alternative imputation strategies.
Section 1.2.1	When a treatment policy strategy is applied for ST initiation ICE, updated to use data irrespective of whether the participant complied with the 12-hour ST washout.	See rationale for update in Section 1.1.

Section #	Description of Change	Brief Rationale
	Statement removed that ANCOVA will be for the ST de-mediated approach at Month 12.	To align with updates in Section 5.
Section 1.2	Table 1-2 updated to note that the Part II emerging symptoms estimand is not covered in this table, and will instead be covered in Section 5.3.2.1.7. Table 1-2 updated to remove footnote [b] which covered ST washout compliance.	As the handling strategies for this estimand did not easily align with the other endpoints, updated to cover this in the specific section for the estimand. See rationale for update in Section 1.1.
Section 3	Where type I error is mentioned, "(1-sided)" has been added.	To clarify that this is the 1-sided type I error.
Section 4	For the Full Analysis Set, DaT-SPECT condition added to the list of inclusion/exclusion criteria to check.	To align with Appendix 6.6.
Section 5.1.1	AE duration calculation updated for partially missing dates.	To ensure that when some information is available for the event data (year or year and month), that information is used in the duration calculation.
Section 5.2	An additional listing has been added to this section for participants excluded from the FAS. A separate summary will also be produced summarizing the reasons for ineligibility for those who screen fail due to being ineligible. Two additional outputs included for use in the study Plain Language Summary.	To produce a listing that clearly identifies those excluded from the FAS and the reason for their exclusion. To allow the most commonly failed eligibility criteria to be easily identified. To ensure all PLS requirements are covered in

Section #	Description of Change	Brief Rationale
		the SAP, these outputs are not required for the CSR.
Section 5.3	In the estimation considerations and convergence issues section, the second bullet point has been updated to add clarity around which models this is applicable for. In the statistical outputs section, paragraph added to describe forest plots that will be produced for some key analyses.	To add clarity when there are convergence issues. To help interpretation of results.
Section 5.3.1	One additional outputs included for use in the study Plain Language Summary, a repeat of the primary analysis table produced to 1 decimal place only.	To ensure all PLS requirements are covered in the SAP, this table is not required for the CSR.
Section 5.3.1.1	Descriptive longitudinal mean plot updated to present both observed mean and mean change from baseline.	To aid interpretation of the final results.
Section 5.3.1.3	Sensitivity analysis added for the primary estimand, adjusting for additional prognostic covariates (e.g., Parkinsonian age/digital twin).	To explore the possibility of including prognostic covariate adjustment for internal decision making. The results will be part of a separate report and may not be part of the CSR.
Sections 5.3.1.4, 5.3.2.1.1, 5.3.2.1.2, 5.3.2.1.3, 5.3.2.1.6 and 5.3.2.1.8	Categorical time LMEMs updated throughout all of these sections, to use a heterogenous auto-regressive (ARH1) variance-covariance matrix to account for the repeated measures within subject in place of a random slope model. In sections 5.3.1.4.2 and 5.3.2.1.2, LMEM plots have been added to the text describing the TFLs to be produced.	To make use of a more intuitive model. To align with the TFLs already planned for production.

Section #	Description of Change	Brief Rationale
Section 5.3.1.4.1	Text updated to align with the methodology to be implemented for the final analyses.	To align with the planned final analysis approach.
Section 5.3.1.4.2	Section updated to remove the ST washout compliance requirement. All post-ST data will be used in the analysis, regardless of washout compliance.	See rationale for update in Section 1.1.
Section 5.3.1.4.3	New section added for an MDS-UPDRS Part I-III sum score supplementary analysis in the absence of ST initiation (de-mediated difference in slope over 12 months).	A higher than anticipated proportion of participants initiated ST prior to Month 12, potentially introducing bias into a hypothetical strategy where post-ST outcome data is censored. Therefore an additional hypothetical strategy where post-ST outcome data is de-mediated has been included.
Section 5.3.2.1.2	Sensitivity analysis added for the primary estimand adjusting for additional prognostic covariates (e.g., Parkinsonian age/digital twin).	To explore the possibility of including prognostic covariate adjustment for internal decision making. The results will be part of a separate report and may not be part of the CSR.
Section 5.3.2.1.4	Additional details for this estimand added to the SAP.	To align with the planned final analysis approach.
Section 5.3.2.1.5	Main analytical approach text updated to clarify that post-study termination data will be right-censored (not ignored).	To align with the ICE approaches outlined in Table 1.2.
Section 5.3.2.1.7	Section updated to include additional details on the ICE handling strategies for this estimand, for both the main and supplementary analyses.	As the handling strategies for this estimand did not easily align with the other endpoints, updated to cover this in the specific section for the estimand. For the alternative Part II endpoint, updates made to align with

Section #	Description of Change	Brief Rationale
		the ICE approaches outlined in Table 1.2.
Section 5.3.2.1.8	Analyses regardless of ST initiation relabelled as 'other secondary' analyses rather than 'supplementary'.	To align with the study protocol.
Section 5.3.2.3.1	Descriptive longitudinal mean plot updated to present both observed mean and mean change from baseline.	To aid interpretation of the final results.
Section 5.4.2	Three additional outputs included for use in the study Plain Language Summary.	To ensure all PLS requirements are covered in the SAP, these outputs are not required for the CSR.
Section 5.4.2.1	Section updated to state that treatment-emergent AESIs will be summarized.	To ensure that these summary tables on present treatment-emergent events.
Section 5.4.3.1.1	Four additional TFLs added for liver function tests. A note was added that the figure for participants who meet the Hy's Law Criteria only needs to be produced if there are participants who meet the criteria.	To allow for a full assessment of hepatic safety in a large Phase 2 study.
Section 5.5.1	Rule added for predose concentration data that is confirmed to not have been collected prior to dosing.	This data should be listed only and not used for descriptive statistics.
Section 6.2	Additional supplementary analyses added in the SAP that are not covered in the protocol, for MDS-UPDRS Part I-III sum score and MDS-UPDRS Part III subscale/subscore ST de-mediated data. Bullet point for MDS-UPDRS Part I at Month 12 and Month 18 regardless of ST initiation being considered supplementary in the SAP removed, these are now	To document additions to the protocol planned analyses. SAP and protocol are now aligned, considered secondary in both documents.

Section #	Description of Change	Brief Rationale
	considered secondary in both the SAP and protocol.	
Section 6.5	New appendix added.	To add in the statistical details for the ST de-mediation analyses.
Section 7	References from appendix 5 added.	To include references from the new appendix.

Amendment 3 (18 Jun 2024)

Overall Rationale for the Amendment

To align with updates made in protocol amendment 6, and to increase efficiency of final study analyses.

Section #	Description of Change	Brief Rationale
Section 1	Protocol version updated.	To reference the latest version of the protocol, amendment 6.
Section 1.1	Table 1-1 updated to align with Table 9-1 in the protocol.	To align with changes made in protocol amendment 6.
	Footnote added to Table 1-1: "Note: The secondary estimands included in this table for MDS-UPDRS Part I at Month 12 and Month 18 regardless of ST are considered supplementary estimands in this SAP and are covered in Section 5.3.2.1.8."	To highlight a difference between the way the protocol and SAP classify these analyses.
	Footnote in Table 1-1 about ST washout compliance updated to clarify that censoring when a participant does not comply with the washout only applies to analyses of MDS-UPDRS Part III and I-III sum score.	To clarify that this censoring does not apply to all analyses.
Section 1.2	Additional details added to the definition of the hypothetical and composite strategies for handling ICEs.	To provide additional information for the hypothetical strategy, and to refer to both Part II and Part II for the composite strategy.
	Table 1-2 and 1-3 updated to include all population summary measures covered in protocol amendment 6.	To align with protocol amendment 6.

Section #	Description of Change	Brief Rationale
	Footnote in Table 1-2 about ST washout compliance updated to clarify that censoring when a participant does not comply with the washout only applies to analyses of MDS-UPDRS Part III and I-III sum score.	To clarify that this censoring does not apply to all analyses.
	Footnote for washout compliance removed from Table 1-3.	Washout compliance is not relevant for the estimands described in this table.
	Footnote [a] updated in Table 1-2. Previously this footnote stated that additional sensitivity analyses would be performed for some of the primary and secondary endpoints, these sensitivity analyses will now be performed outside of the SAP and the footnote is now used for the Part III ePD subscore.	To align with planned approach for sensitivity analyses.
	Additional footnotes added to Table 1-2 for MDS-UPDRS.	To provide additional information for handling missing data.
	Section on study termination and loss to follow-up updated in Table 1-2 and 1-3 to cover missed visits as well as termination. Handling descriptions updated for this section to not use ICE terminology. Additional details added for the handling of time-to-event/binary endpoints.	To ensure all endpoints are covered, and to provide strategies for handling missed visits.

Section #	Description of Change	Brief Rationale
	A row for scanner changes has been added to Table 1-3.	To confirm that scanner changes will be handled using a treatment policy approach for DaT-SPECT analyses.
	ICE handling strategy descriptions have been updated for event-based outcomes in Table 1-2.	To clarify the handling strategies.
Section 1.2.1	Updates made to ensure the only ICE discussed in this section is ST initiation, not dose or type changes.	Changes in medication type or dosage will not be considered as ICEs in this study.
	For MDS-UPDRS, updated to discuss de-mediation instead of control-based mean imputation.	To align with the estimands defined in protocol amendment 6.
	For DaT-SPECT, removed the discussion of the 'in the absence of ST' analyses.	These analyses have been removed from Section 5 of this SAP. The analyses of DaT-SPECT in the absence of ST initiation will now be performed as part of the exploratory analyses outside of the CSR.
	Paragraph about ST washout compliance updated to clarify that censoring when a participant does not comply with the washout only applies to analyses of MDS-UPDRS Part III and I-III sum score.	To clarify that this censoring does not apply to all analyses.
Section 1.2.5	Title updated to remove 'COVID-19 Vaccination'.	Vaccination is no longer consider as its own separate ICE.

Section #	Description of Change	Brief Rationale
Section 1.2.6	Section updated to align with updates made in Table 1-2 and 1-3.	To align with the ICE handling strategies defined for each estimand.
Section 1.3	Study schematic figure removed, replaced with a reference to protocol section 1.2.	To reduce the risk of inconsistencies if the schema is updated in the protocol.
Section 4	Statement added to each analysis set definition to state whether outputs using each set will use planned or actual treatment assignment. FAS definition updated to exclude participants who do not meet key inclusion/exclusion criteria. Statement added that for all analyses, a participant should only be included in the model if they have a baseline score for that endpoint.	To add in additional details required for TFL programming. To ensure our FAS only includes participants who are part of our target population. To ensure comparability between main, sensitivity and supplementary analyses for the same endpoint.
Section 5.1	Statement added that some analyses may not be performed by Parexel. Paragraph discussing the Month 12 analyses removed.	The new estimand for MDS-UPDRS Part III added during this SAP amendment will be performed in-house by the UCB Statistical team. To align with protocol amendment 6. All analyses over 12 months will performed at the end of the study, not at the time where all participants reach Month 12.
Section 5.1.2.3	Data display labels updated to use ‘Minzasolmin’ in place of ‘UCB0599’.	To align with other study documents.

Section #	Description of Change	Brief Rationale
Section 5.1.2.6	Summary tables by country and region updated to be by region only. Additional endpoints added to the list to be summarized by region: MDS-UPDRS Part III ePD subscore, MDS-UPDRS Part III first 5-point increase and MDS-UPDRS Part II first 3-point increase.	To focus on key trends and avoid summaries based on small numbers of participants. To include summaries by region for some additional key endpoints.
Section 5.3	The following was added for the mean centering of covariates: "When mean centering covariates, the calculated mean will be based on the data used for that estimand analysis." Statement added that for all analyses, a participant should only be included in the model if they have a baseline score for that endpoint. Disease duration removed as a covariate to be included in all efficacy models. Section defining the MDS-UPDRS sum scores updated to include the range for each score.	To clarify how to perform the mean centering. To ensure comparability between main, sensitivity and supplementary analyses for the same endpoint. Disease duration is no longer considered a reliable predictor of disease progression, and could make results interpretation more complex. To provide additional information for each sum score.

Section #	Description of Change	Brief Rationale
	ST initiation timing categories for figures updated so that the 'not started before Month 18 category' included Month 18 or EOT, so that early terminators are included in this category.	Clarify how early terminators should be included in these figures.
	Text defining the baseline predictors of ST initiation removed.	All sensitivity analyses adjusting for these predictors have been removed as it is unlikely to account for the strong bias introduced due to the large proportion of participants initiating ST prior to Month 12 in analyses where post-ST initiation data is censored. Such sensitivity analyses may become part of exploratory analyses. Therefore, the definition is no longer required.
	Section added for handling convergence issues.	To provide a detailed description of how to handle these issues if they do occur.
	Statement added to the statistical outputs section that for some key analyses, the summary table will also present the percentage decrease in mean slope or mean difference at Month 18 for each dose-level against placebo.	To ensure that the statistical outputs include all required information for results interpretation and decision-making.

Section #	Description of Change	Brief Rationale
Section 5.3.1	MDS-UPDRS descriptive summary statistic tables updated to use a different definition of ST status. Previously ST status was based on status at Month 12 and 18, in this amendment we are updating to summarize by status at each individual visit. Summary tables of mean MDS-UPDRS scores over time updated to not be produced by gender.	To align with the summary tables for all other endpoints, where this approach is already being used. To simplify these summary tables.
Section 5.3.1.2	Statement added that time in the mixed effects model should be time since baseline. Section on handling convergence issues removed.	To provide additional details for the programming team. This is now included in section 5.3 and applies to all mixed effects models.
Section 5.3.1.3	All sensitivity analyses for the primary estimand analysis removed.	The primary estimand with its hypothetical strategy for rescue medication ICE handling where post-ST initiation data is censored is no longer considered sensitive and unbiased due to the high proportion of participants who initiated ST prior to Month 12 and the impact of ST of the Part III subscale. Therefore, sensitivity analyses to the main estimator are no longer required. Instead, a supplementary analysis is proposed where post-ST initiation MDS-UPDRS Part I-III data is de-mediated and the difference

Section #	Description of Change	Brief Rationale
		between treatment arms is analysed at Month 12.
Section 5.3.1.4	Split into two sections: 5.3.1.4.1 and 5.3.1.4.2. 5.3.1.4.1 includes a new supplementary analysis for this endpoint that will be performed in-house by UCB. 5.3.4.1.2 describes the original supplementary analysis planned for this estimand.	To include a new supplementary analysis for the primary estimand,
Section 5.3.2.1,	Sentence added to the section describing the secondary estimands: "In addition, a sub-score based on the MDS-UPDRS Part III and targeted at the early-stage PD population will also be analysed."	To describe a new endpoint added in this amendment.
Sections 5.3.2.1.1, 5.3.2.1.2 and 5.3.2.1.3	All sensitivity analyses removed for the following secondary estimands: • MDS-UPDRS Part I-III sum score at 18 months (in the absence of ST initiation, with censoring) • MDS-UPDRS Part III subscale slope of progression over 12 months (in the absence of ST initiation), • MDS-UPDRS Part III subscale at 18 months (in the absence of ST initiation).	The same rationale applies here as for the primary estimand.
Section 5.3.2.1.4	New MDS-UPDRS Part III subscale and MDS-UPDRS Part III ePD subscore estimands added: ST de-mediated estimands, i.e. as if the participants had not initiated ST	An estimand based on the MDS-UPDRS Part III subscale and a reduced version of the subscale more targeted at the early-stage population (ePD subscore) has been added where the impact of ST is removed from the data collected after the initiation of rescue

Section #	Description of Change	Brief Rationale
		medication using de-mediation approaches.
Section 5.3.2.1.5 (previously Section 5.3.2.1.4)	Definition of time to worsening updated to be based on a *confirmed* 5-point increase in MDS-UPDRS Part III. ICE handling strategies updated. Analyses of time to worsening within 12 months removed. Summary table for this endpoint has been updated to not be produced by gender and to align with the new definition of time to worsening.	To harmonize our approach for MDS-UPDRS time-to-event endpoints. To align with Table 1.2. To align with protocol amendment 6 where this estimand has been removed. To simplify this summary table.
Section 5.3.2.1.6 (previously Section 5.3.2.1.5)	Analysis at 12 months removed. Sensitivity analysis adjusting for baseline predictors of ST initiation removed. For analyses using log-transformed data, a rule has been added for scores of 0.	To align with protocol amendment 6, estimates at 12 months will come from the analysis at 18 months. The same rationale applies here as for the primary estimand. To provide additional details for data handling.
Section 5.3.2.1.7	New MDS-UPDRS Part II subscale-based estimand added.	To add a new secondary efficacy estimand based on MDS-UPDRS Part II data already collected in the study.

Section #	Description of Change	Brief Rationale
Section 5.3.2.1.8 (previously Section 5.3.2.1.6)	Analysis at 12 months removed. Main analysis approach updated to be in the absence of ST. Supplementary analysis added for this estimand regardless of ST. For analyses using log-transformed data, a rule has been added for scores of 0.	To align with protocol amendment 6, estimates at 12 months will come from the analysis at 18 months. To align with protocol amendment 6 where the estimand in the absence of ST is the main analysis. To provide additional details for data handling.
Section 5.3.2.2.1	Descriptive statistics summary table for MoCA updated to not be produced by gender.	To simplify these summary tables.
Section 5.3.2.2.2	Note added that this model should be fit using the SAS PROC MIXED procedure.	To provide additional details for the programming team.
Section 5.3.2.2.3	Section describing sensitivity analyses for this estimand removed.	In order to increase efficiency of final study analysis performance and review, any sensitivity analyses required for this endpoint may now be performed as part of exploratory analyses.
Section 5.3.2.3.1	Descriptive statistics summary table for DaT-SPECT updated to not be produced by gender.	To simplify these summary tables.
Section 5.3.2.3.2	ANCOVA for Month 12 data removed.	Estimates for Month 12 will come from the model fitted to the data from baseline to Month 18.

Section #	Description of Change	Brief Rationale
Section 5.3.2.3.3	Sensitivity analyses adjusting for baseline predictors of ST initiation removed. Control-based mean imputation sensitivity analysis removed. The previous sensitivity analysis using the original data has been replaced with two analyses. One where baseline is part of the responses and time is continuous, and one where baseline is used as a covariate.	In order to increase efficiency of final study analysis performance and review, any sensitivity analyses required for this endpoint may now be performed as part of exploratory analyses. To add an extra sensitivity analysis which makes use of the original data.
Section 5.3.2.3.4	Supplementary analysis in the absence of ST removed.	A hypothetical strategy for rescue medication ICE handling where post-ST initiation data is censored is no longer considered sensitive and unbiased due to the high proportion of participants who initiated ST prior to Month 12. In addition, it is believed that DaT-SPECT signal is not affected by ST initiation. Therefore, this supplementary analysis will now be considered an exploratory analysis.
Section 5.3.2.4.1	ICE handling strategy statement updated to refer to Table 1.3 rather than describe the strategies in text.	To avoid duplication.
Section 5.3.2.4.2	ICE handling strategy statement updated to refer to Table 1.3 rather than describe the strategies in text.	To avoid duplication.

Section #	Description of Change	Brief Rationale
	Sensitivity analyses adjusting for baseline predictors of ST initiation removed.	In order to increase efficiency of final study analysis performance and review, any sensitivity analyses required for this endpoint may now be performed as part of exploratory analyses.
Section 5.3.3.1	Summary tables for time to first increase in MDS-UPDRS Part I/II specific items removed.	In order to increase efficiency of final study analysis performance and review, these analyses may now be performed as part of exploratory analyses.
	Listing of MDS-NMS non-motor fluctuations subscale data removed.	Subscale removed from the TFLs, it is the main MDS-NMS questionnaire that is of interest in PD0053.
	Summary tables for exploratory endpoints updated to remove summarizing by gender.	To simplify these tables.
	MDS-UPDRS selected items section removed and replaced with a section for the exploratory composite endpoint (based on MDS-UPDRS and/or Early PD PROs).	To align with the exploratory endpoints defined in protocol amendment 6, these analyses will be covered as part of an exploratory analysis plan.
Section 5.3.3.2	Summary tables for these exploratory endpoints updated to remove summarizing by gender.	To simplify these tables.
Section 5.4.2.1	Statement added to note that AESIs will also be identified by the investigator, the investigator flag will be used for flagging events in	To clarify how the AESI listings and summary tables are produced.

Section #	Description of Change	Brief Rationale
	our listings and for presenting summary tables. An additional summary table has been added to this section, summarizing AESIs by relationship. A separate listing of AESIs has also been added.	To allow the team to easily identify the number of AESIs and treatment-related AESIs.
Section 5.6	Sensitivity analyses by gender removed for: MDS-UPDRS Part II, III, I-III sum score and ST intake at 12 months. Sentence added: "Key analyses may be repeated among participants who did not experience tremor at baseline as part of the exploratory analysis plan."	In order to increase efficiency of final study analysis performance and review, these analyses may now be performed as part of exploratory analyses. To describe additional analyses that may be performed outside of the main analysis plan.
Section 5.8	Section aligned with protocol amendment 6. The first paragraph was revised to clarify that the analyses of Month 12 data will be carried out at study end. A paragraph was added to describe that a limited team of independent pharmacometricians and statistical programmers with access to the 12-month data will be unblinded to prepare for the exposure-response analysis.	To clarify that no primary analysis at Month 12 will be done, and to describe the unblinding of an independent expert group to develop the exposure-response model.
Section 6.1.1	Disease duration calculation updated to be the time between diagnosis and first dose of IMP, rather than the time between diagnosis and ICF.	To align with all other baseline characteristics where data collected at the baseline visit is presented.
Section 6.2, Appendix 2	All text from SAP amendment 2 removed. Section updated to cover	SAP amendment 2 updates that deviated from the

Section #	Description of Change	Brief Rationale
	additions in SAP amendment 3 that are not covered in protocol amendment 6.	protocol have been added to protocol amendment 6.
Section 6.5, Appendix 5	New section added.	To provide additional details on the imputation and de-mediation approaches. Final details for these approaches will be provided as part of the last SAP amendment before study readout.
Section 6.6, Appendix 6	New section added.	To provide criteria for excluding participants from the FAS.
Section 7, References	New references added: • Floden and Bell, 2019 – referenced in section 1.2 • Lasch et al., 2022 – referenced in section 1.2 • Brown and Prescott, 2014 – referenced in section 5.3	New references included for using the hypothetical estimands strategy and handling convergence issues.

Amendment 2 (5 October 2023)

Overall Rationale for the Amendment

To add additional details on data handling rules.

Section # and Name	Description of Change	Brief Rationale
Global	Updates made throughout to change 'Compliance' to 'IMP Compliance' where applicable.	As we also discuss ST washout compliance in this SAP, text updated to make it clear which compliance we are referring to.
Section 1.2	In the tables of ICEs: "Important ICE-like protocol deviations related to investigational treatment" updated to "Important ICE-like protocol deviations". Footnotes updated to add details on the assumptions when determining whether participants were compliant with the ST washout.	Protocol deviation specifications document is now the main source for selecting which important protocol deviations are ICE-like. To help determine compliance when not all dates and times are captured.
Section 1.2.4	The following sentence was added to this section: "In practice, these events will be identified as serious adverse events that result in persistent or significant disability/incapacity or death.".	To help identify these events.
Section 1.2.6	Statement added to state that "Study termination will only be considered treatment-related if the reason for termination is 'lack of efficacy'.".	To help identify treatment-related study termination events.
Section 5.1.1	Rules for calculating duration of Parkinson's Disease updated to use the middle of the month/year rather than the beginning of the month/year when dates are partial.	Update to our analysis approach, assume a central date rather than assuming that the participant had the longest possible disease duration.

Section # and Name	Description of Change	Brief Rationale
	Rule added for determining ST initiation/end date to be used in analyses when partial dates are captured.	Add in the assumption to be made when day is missing from ST initiation/end date, so that ICE strategies can be correctly applied in these cases and LEDD can be calculated.
Section 5.1.2.1.4	Text added to state that DaT-SPECT data will use different mapping rule to other endpoints. Reference added to Section 5.3.2.3 for these rules.	As scans are only planned three times throughout the study, mapping rules for assessments performed more regularly do not apply.
Section 5.3.2.3	Rules for mapping DaT-SPECT data added.	See row above for rationale.
Section 6.1.1.2	MDS-UPDRS, MoCA and Hoehn and Yahr updated to be summarized based on Day 1 data, or the latest available pre-baseline data if no Day 1 is available.	To allow the latest available data to be used in baseline summaries.
Section 6.1.8	Listing and summary of compliance data updated to also present compliance calculated under the assumption that no overdosing has occurred.	Additional compliance calculations were added that assume any discrepancies between planned and actual dosing are due to drug accountability issues and do not necessarily represent participants taking more than the planned doses.
Section 6.4	Table title updated from 'Overview of handling strategies for ICEs and study termination for other estimands' to 'Treatment groups to be presented in summary tables'.	Previous title was incorrect.

Amendment 1 (25 Jan 2023)

Overall Rationale for the Amendment

Align with current Estimands thinking and add additional details on data handling rules based on discussions within the stats and programming teams.

Section # and Name	Description of Change	Brief Rationale
Global	References to a Month 12 SAP removed throughout.	A charter will be produced instead of an analysis plan.
Section 1.2	Moved ICE strategy definitions from Section 5.1.1 to Section 1.2. Updated strategy for handling ICE-like protocol deviations. ICE language updates made throughout.	ICE information all in one section now for clarity, improvements made to protocol deviation handling and ICE language to best-align with the current Estimands guidelines.
Section 5.1	Text added for handling unscheduled visit data, this data will be listed but not included in most summary tables.	Add clarity for the programming team.
Section 5.1.1	Imputation rules for calculating duration of Parkinson's Disease added.	Missing from last version of the SAP, rules for AEs and concomitant medications were not applicable to disease duration.
Section 5.1.2.1.4	Text added for handling early safety follow-up visit data.	Add clarity for the programming team.
Section 5.1.2.6	Summary of number of participants who have received the COVID-19 Vaccine removed.	Less focus needed on COVID-19 and vaccination status, a large number of participants expected to be vaccinated.
Section 5.3	Disease duration added as a covariate in all efficacy models. Detail added on handling model convergence issues. MDS-UPDRS sum score calculation details added.	Improve the modelling strategy. Add details that aren't in questionnaire manuals.

Section # and Name	Description of Change	Brief Rationale
	ST status definition added.	Add details of assumptions that will be made for ST handling.
Section 5.3.2.2	Added text describing how to handle 'years of schooling' information from the MoCA questionnaire.	Information not included in the questionnaire manual.
Section 5.3.2.3 and 5.3.3.2	Updates made to refer to DaT-SPECT 'whole' striatum not 'mean' striatum throughout.	To maintain consistency.
Section 5.3.3.3	Levodopa equivalent daily dose conversion factors removed, reference added to ADRG instead.	As this will be a live document that needs to be fed into the datasets, removed from the analysis plan.
Section 5.4.2	Two mandatory AE tables added: Incidence of serious TEAEs by relationship and Incidence of fatal TEAEs by relationship.	To comply with mandatory reporting requirements.
Section 5.4.3.1	Text added for handling clinical safety laboratory retest results.	Add clarity for the programming team.
Section 5.4.3.6	Listing of MRI dates removed.	Information will be available in the database, this listing would not add value.
Section 5.9	PK removed from the list of data the DMC will review.	To align with current DMC charter, PK data will not be reviewed as part of the study safety assessments.
Section 6.1.4	Concomitant medication definition updated to extend beyond date of last dose of study medication.	To align with the half-life of UCB0599, drug will still be in a participants system several days after the last dose.
Throughout	Minor editorial and document formatting revisions.	Minor, therefore have not been summarized.

LIST OF ABBREVIATIONS

AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase
ANCOVA	analysis of covariance
ASPS	All Study Participants Screened
AST	aspartate transaminase
ASYN	alpha-synuclein
BID	twice per day
BLQ	below limit of quantification
BMI	body mass index
BP	blood pressure
CGII	Clinical Global Impression of Improvement
CGIS	Clinical Global Impression of Severity
CI	confidence interval
COVID-19	Coronavirus Disease-2019
CRO	Contract Research Organization
CSF	cerebrospinal fluid
CSR	Clinical Study Report
C-SSRS	Columbia-Suicide Severity Rating Scale
CV	coefficient of variation
DaT-SPECT	Dopamine Transporter Imaging with Single Photon Emission Computed
DEM	data evaluation meeting
DM	disease modification
DMC	Data Monitoring Committee
DMT	disease modifying therapy
DNA	deoxyribonucleic acid
DRM	data review meeting
ECG	electrocardiogram
eCRF	electronic Case Report Form
EMA	European Medicines Agency
EOT	End of Treatment

EQ-5D-5L	Euro Quality of life 5-Dimensions 5-Level
EQ-VAS	EQ visual analogue scale
EudraCT	European Union Drug Regulating Authorities Clinical Trials
FAS	Full Analysis Set
FDA	US Food and Drug Administration
FSH	follicle stimulating hormone
GCP	Good Clinical Practice
HADS	Hospital Anxiety and Depression Scale
HLT	high level term
HR	hazard ratio
HSR	hypersensitivity reaction
ICF	Informed Consent Form
ICH	The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICE	intercurrent event
IMP	investigational medicinal product
IPD	important protocol deviation
IQR	Inter quartile range
IRB	Institutional Review Board
IRT	interactive response technology
LLOQ	lower limit of quantification
LMEM	linear mixed effects model
LOCF	last observation carried forward
MAR	missing at random
MCMC	Markov Chain Monte Carlo
MDS-NMS	Movement Disorder Society Non-motor symptom scale
MDS-UPDRS	Movement Disorder Society Unified Parkinson's Disease Rating Scale
MedDRA	Medical Dictionary for Regulatory Activities
MEM	mixed effects model
MI	multiple imputation
MNAR	missing not at random
MoCA	Montreal Cognitive Assessments
MRI	magnetic resonance imaging

NDD	neurodegenerative disease
PD	Parkinson's Disease
PDILI	potential-drug induced liver injury
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	pharmacokinetic(s)
PKS	Pharmacokinetic Set
POC	proof of concept
PPMI	Parkinson's Progression Markers Initiative
PRO	patient reported outcome
PT	preferred term
QTcF	QT corrected for heart rate using Fridericia's formula
RBI	Reference-Based Imputation
RMET	restricted (t-year) mean event time
RNA	ribonucleic acid
RS	Randomized Set
SAE	serious adverse event
SAP	statistical analysis plan
SARS-CoV-2	severe acute respiratory syndrome coronavirus-2
SBR	specific binding ratio
SD	standard deviation
SE-ADL	Schwab and England Activities of Daily Living
SFU	Safety Follow-up
SMC	Safety Monitoring Committee
SMQ	Standardized MedDRA Query
SNMM	structural nested mean model
SOC	system organ class
SS	Safety Set
ST	symptomatic treatment
SWEDD	scans without evidence for dopaminergic deficit
TEAE	treatment-emergent adverse event
TFL	tables, figures and listings
ULN	upper limit of normal

VME	Virtual Motor Exam
WHODD	World Health Organization Drug Dictionary

1 INTRODUCTION

The purpose of this statistical analysis plan (SAP) is to provide all information that is necessary to perform the required statistical analysis of study PD0053. It also defines the summary tables, figures and listings (TFLs) to be included in the final clinical study report (CSR) according to the protocol.

This SAP is based on protocol amendment 6, dated 21 March 2024.

Unless specified in the sections below, the study will be analyzed as described in the most recent version of the protocol. If a future protocol amendment necessitates a substantial change to the statistical analysis of the study data, this SAP will be amended accordingly. In addition, if analysis definitions must be modified or updated prior to database lock, a SAP amendment will be required. If, after database lock, additional analyses are required to supplement the planned analyses described in this SAP, these changes will be described in the CSR together with the associated rationale. Other minor changes to non-key analyses will also be documented in the CSR.

The content of this SAP is compatible with the International Council for Harmonization (ICH)/Food and Drug Administration (FDA) E9 Guidance documents (Phillips et al, 2003).

UCB is the sponsor and Parexel is the Contract Research Organization (CRO) for this study.

1.1 Objectives and estimands/endpoints

The main objective of this study is to provide proof of concept (POC) for the efficacy of the alpha-synuclein (ASYN) misfolding inhibitor UCB0599 in reducing disease progression in study participants with early-stage Parkinson's disease (PD), and to instruct later stage development. The ultimate goal is to provide novel treatment options to PD patients which have the potential to modify the progression of the disease.

Two types of data have been identified as possible supportive evidence of Disease Modification Therapy (DMT) (Cummings, 2017):

(1) The intervention produces a significant drug-placebo difference on accepted clinical outcome(s), AND has a consistent effect on (fluid or imaging) biomarkers considered reflective of the fundamental pathophysiology of a Neurodegenerative Disease (NDD), ie, neurodegeneration.

(2) The intervention produces a positive outcome on a randomized start clinical study design consistent with an enduring (positive) change (ie, treatment effect) in the clinical course of disease.

However, the use of randomized start designs to show disease modification remains controversial.

Therefore, the proposed study, PD0053, was chosen to be a Phase 2a study with a double-blind parallel-group, placebo-controlled design and a treatment duration of 18 months.

Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Parts I-III sum score, the primary clinical outcome, is the best-established functional scale to evaluate disease progression in PD populations and to assess the effect of symptomatic treatments (STs). ST currently available for PD patients, namely levodopa and other dopamine agonists, will interfere with motor symptoms and therefore bias any MDS-UPDRS measures collected post-initiation of ST. The intention with this study is to provide a measure of the effect of UCB0599 on the MDS-UPDRS clinical scale, free of confounding by ST.

In the case of drugs with the potential of slowing disease progression in this population, MDS-UPDRS would have considerable limitations if applied as a solitary and standalone endpoint. As it is currently the best characterized and most widely used scale, MDS-UPDRS will have a prominent role in this study to assess the effect of UCB0599 in delaying the progression of clinical symptoms; however, it is understood that it almost exclusively captures neuromotor dysfunction and thus in isolation would not have the bandwidth to provide a holistic picture of disease progression. To address this limitation and to ensure a more holistic view, the suggested POC study incorporates multiple measures and scales capturing non-motor domains (including cognition and neurovegetative function).

Evaluation of how study participants feel and function through the use of patient-reported outcome (PRO) measures has become an increasingly important component of therapeutic assessment. In an effort to best measure participant perception of motor and non-motor signs and their impact in early-stage PD, UCB, together with the patient organizations Parkinson's UK and Parkinson Foundation USA, has collaborated on large patient-centered research to define an optimized PRO strategy in the context of use for early-stage PD. Research findings across over 50 patients and 10 caregivers highlighted that bradykinesia (particularly function slowness), tremor, rigidity/stiffness, mobility (particularly gait and upper limb issues), and fatigue are the cardinal concepts of interest in early-stage PD. Research outcomes also suggested that the legacy PROs traditionally used in clinical research may not be fit for purpose in early-stage PD, as most of them were developed for use in later stages of disease. As a result, UCB developed with patients an exploratory set of novel PROs for early-stage PD, composed of: the Early Parkinson's Disease Function Slowness PRO, the Early Parkinson's Disease Mobility PRO, and Fatigue-PRO. UCB hypothesizes that this exploratory PRO strategy will generate interpretable PRO data for the demonstration of treatment benefit in early-stage PD patients and adequately complement MDS-UPDRS evidence. It is intended to test the new UCB PRO strategies in PD0053 as exploratory endpoints.

To assess neurodegeneration, the biomarker Dopamine Transporter Imaging with Single Photon Emission Computed Tomography (DaT-SPECT) will be used, which is the best-established in vivo method to monitor dopaminergic neurodegeneration. It has been reported in the literature that DaT-SPECT results are not impacted (or are only marginally impacted) by using ST (Ikeda et al, 2019). Therefore, DaT-SPECT will be used as an assessment in the study to identify the effect of UCB0599 treatment. Moreover, DaT-SPECT at Screening will allow identification of the relevant target population by identifying patients without evidence of dopaminergic deficit.

Previous studies that have attempted to demonstrate efficacy of non-symptomatic treatments for PD in early-stage populations have failed to show disease modification. Among the reasons for failure, insufficient treatment duration to detect delayed treatment effect and the high variability of disease progression have been identified. An 18-month duration was chosen because it is expected to be sufficient time for UCB0599 to reveal a disease modifying effect and to differentiate from ST.

In addition, the impact of UCB0599 on time-to-initiation and overall intake of ST will be investigated. Since UCB0599 treatment is expected to slow disease progression and disease progression is directly linked to onset of ST, UCB0599 would be expected to delay the time until the study participants require initiating ST, which therefore can be utilized as an event. The treatment duration of up to 18 months means that treatment effects of UCB0599 on the need for initiation of ST will be able to be assessed in the majority of study participants.

As UCB0599 has an additional potential to slow disease progression in the non-motor domain, which prominently encompasses cognitive functioning, there is a need to include instruments which may capture a signal in this area. The Montreal Cognitive Assessments ([MoCA]; Nasreddine et al, 2005) will be used to complete the main assessments, as it covers different cognitive domains. It is available in parallel versions (with only limited training effects and well-established validity in the PD population) and can thus be utilized for longitudinal observations. Furthermore, the MoCA will play a role in Screening, to exclude participants with a pronounced cognitive impairment, which specifically in younger participants could be indicative of a neurodegenerative disorder other than PD.

For each efficacy objective, the precise treatment effect of interest is defined on the selected endpoints referred to as the estimands. As this is a Phase 2 study, we will focus on estimands which measure efficacy (as opposed to effectiveness).

In the study protocol, an objectives and endpoints table is presented in section 3 and a version presenting estimands for the efficacy objectives is presented in section 9.3. The table below is a combination of the two, using the estimands framework from section 9.3 of the protocol, but also including the safety and PK objectives from the objectives and endpoints table. Other secondary estimands have also been included in this table which were not included in the protocol objectives tables. Supplementary estimands and sensitivity analyses are not covered in this table, these will be discussed in Section 5.3.

For all estimands, the target population will be the entire study population.

For all estimands, the difference between UCB0599 and placebo will be evaluated for both the low and the high doses of UCB0599 (180mg/day and 360mg/day).

In all analyses, the 2 dose levels will be assessed separately against placebo, there will be no pooling of the 2 dose levels, and they will not be compared to each other.

The comparison of the high dose of UCB0599 with placebo will be considered as the primary evaluation. If the comparison of the low dose and placebo gives a positive result on the primary estimand but the comparison of the high dose and placebo does not, these results will have to be interpreted with caution.

Table 1–1: Study objectives

Objectives	Estimands
Primary Efficacy Objective	
To demonstrate the superiority of UCB0599 over placebo with regard to **clinical symptoms of disease progression** over 12 and 18 months in participants diagnosed with early-stage PD	**Primary Efficacy Estimand** • Difference between UCB0599 and placebo in target population mean slope of progression in **MDS-UPDRS Parts I-III sum score** over 12 months in the absence of concomitant ST intake **Key Secondary Efficacy Estimands** • Difference between UCB0599 and placebo in target population mean in **MDS-UPDRS Parts I-III sum score** at 18 months in the absence of concomitant ST intake • Difference between UCB0599 and placebo in target population mean slope of progression in **MDS-UPDRS Part III subscale** score over 12 months in the absence of concomitant ST intake • Difference between UCB0599 and placebo in target population mean in **MDS-UPDRS Part III ePD subscore** at 12 months in the absence of concomitant ST intake • Difference between UCB0599 and placebo in target population RMET, in this case, time to worsening of the disease as defined by a **5-point increase in MDS-UPDRS III**, within the 18-month period, in the absence of concomitant ST intake **Secondary Efficacy Estimands** • Difference between UCB0599 and placebo in target population mean **MDS-UPDRS Part I/II subscales** at 12 months in the absence of concomitant ST intake • Difference between UCB0599 and placebo in target population mean **MDS-UPDRS Part I/II/III subscales** at 18 months in the absence of concomitant ST intake

Table 1–1: Study objectives

Objectives	Estimands
	• Difference between UCB0599 and placebo in target population mean in **MDS-UPDRS Part III ePD subscore** at 18 months in the absence of concomitant ST intake • Ratio between UCB0599 and placebo based on population annualized rate of emerging symptoms assessed by **MDS-UPDRS Part II subscale** over the 18-month period • Difference between UCB0599 and placebo in target population mean **MDS-UPDRS Part I subscale** at 12 months regardless of concomitant ST intake • Difference between UCB0599 and placebo in target population mean **MDS-UPDRS Part I subscale** at 18 months regardless of concomitant ST intake • Difference between UCB0599 and placebo in target population mean observed **MoCA** at 18 months, regardless of concomitant ST intake **Exploratory Efficacy Endpoints** • Time to worsening on MDS-UPDRS Part I subscale • Modified Hoehn and Yahr staging • CGII • CGIS • Fatigue-PRO • Early PD Function Slowness PRO • Early PD Mobility PRO • Composite endpoint (based on MDS-UPDRS and/or Early PD PROs) • PGIS, overall and fatigue-specific • PGIC, overall and fatigue-specific • SE-ADL • HADS • MDS-NMS • Starkstein Apathy Scale

Table 1–1: Study objectives

Objectives	Estimands
	• EQ-5D-5L • Wearable sensor
Secondary Efficacy Objectives	
To demonstrate the superiority of UCB0599 over placebo with regard to **neurodegeneration of dopaminergic neurons** over 12 and 18 months in participants diagnosed with early-stage PD	**Secondary Efficacy Estimand** • Difference between UCB0599 and placebo in target population mean change from Baseline (Screening) in **DaT-SPECT mean Striatum SBR** at 18 months regardless of concomitant ST intake • Difference between UCB0599 and placebo in target population mean change from Baseline (Screening) in **DaT-SPECT mean Striatum SBR** at 12 months regardless of concomitant ST intake **Exploratory Efficacy Endpoints** • DaT-SPECT mean SBR in striatal subregions • CSF total ASYN • CSF ASYN oligomers/seeding capacity
To assess the effect of UCB0599 vs placebo with regard to **intake of ST** over 18 months in participants diagnosed with early-stage PD	**Secondary Efficacy Estimands** • Difference between UCB0599 and placebo in target population RMET, in this case, **time to start of ST**, within the 18-month period, regardless of adherence to assigned study medication • Target population odds ratio between UCB0599 and placebo in **ST intake** at 18 months, regardless of adherence to assigned study medication **Other Secondary Efficacy Estimand** • Target population odds ratio between UCB0599 and placebo in **ST intake** at 12 months, regardless of adherence to assigned study medication **Exploratory Efficacy Endpoints** • Levodopa cumulative daily dose

Table 1–1: Study objectives

Objectives	Estimands
Primary Safety Objectives	
To assess the safety and tolerability of UCB0599 in participants diagnosed with early-stage PD	**Secondary Safety Endpoints** • Incidence of TEAEs • Incidence of SAEs • Incidence of TEAEs leading to participant withdrawal **Other Safety Endpoints** • Change from Baseline in clinical laboratory values (chemistry, hematology, and urinalysis) • Change from Baseline in vital signs • Change from Baseline in physical examination • Change from Baseline in neurological examination findings • C-SSRS findings • ECG findings
Exploratory PK Objective	
To assess the PK of UCB0599 and its N-oxide metabolite in participants diagnosed with early-stage PD	**Exploratory PK Endpoint** • UCB0599 and N-oxide metabolite plasma and CSF concentrations

ASYN=alpha-synuclein; C-SSRS = Columbia-Suicide Severity Rating Scale; CGII=Clinical Global Impression of Improvement; CGIS=Clinical Global Impression of Severity; CSF=cerebrospinal fluid; DaT-SPECT=Dopamine Transporter Imaging with Single Photon Emission Computerized Tomography; ECG = electrocardiogram; EQ-5D-5L=Euro Quality of life 5-Dimensions 5-Level; HADS=Hospital Anxiety and Depression Scale; MDS-NMS=Movement Disorder Society-Non-motor Scale; MDS-UPDRS=Movement Disorder Society-Unified Parkinsons Disease Rating Scale; MoCA=Montreal Cognitive Assessment; PD = Parkinson's disease; PGIC=Patient Global Impression of Change; PGIS=Patient Global Impression of Severity; PK = Pharmacokinetic; PRO=Patient-Reported Outcome; RMET=restricted (t-year) mean event time; SAE = Serious Adverse Event; SBR=specific binding ratio; SE-ADL=Schwab and England Activities of Daily Living; ST = Symptomatic Treatment; TEAE = Treatment-emergent Adverse Event

Note: "in the absence of concomitant ST intake" means that all data recorded after a participant initiates ST will be censored. "regardless of concomitant ST intake" means that participants' data will be included in the analyses even after they have initiated ST.

Note: The secondary estimands included in this table for MDS-UPDRS Part I at Month 12 and Month 18 regardless of ST are covered in Section 5.3.2.1.8 under the paragraph entitled "Other secondary estimand".

1.2 Intercurrent events

The tables below cover all planned efficacy analyses, except for sensitivity analyses.

The strategies/approaches for handling ICEs will be as follows:

- *Treatment Policy*: all data will be included in the analysis, regardless of whether the participant remains on the assigned investigational treatment or discontinued. This will reflect the treatment effect regardless of the ICE.
- *Hypothetical*: data will be modified to mirror its value had an ICE not happened (under some hypothetical conditions). One such hypothetical condition is that the ICE occurred completely at random. In this case, all post-ICE data are set to missing (removed) and imputed under the assumption of missing at random (MAR). This strategy can be applied to ICEs which are considered uninformative with respect to the effect of interest. As a default approach, this strategy will be applied to any occurrence of an ICE affecting the existence of measurements and maximum likelihood (ML) imputation will be used for continuous gaussian-distributed endpoints where a linear mixed effect model is used for analysis (Schafer and Graham, 2002). Alternative imputation approaches will also be considered for ICEs which are considered informative with respect to the effect of interest, under the assumption of missing not at random (MNAR), such as reference-based imputation (RBI) approaches (Carpenter, Roger & Kenward, 2013) for continuous endpoints and non-response imputation (NRI) where missing responder status is imputed as non-response for binary endpoints (Floden and Bell, 2019), as well as causal inference methods (Lasch et al., 2022). This approach will reflect the treatment effect had the ICE not occurred.
- *Composite*: ICE data will be incorporated into the endpoint being analyzed. This approach will be used for the MDS-UPDRS Part II/III time to worsening analyses; a composite endpoint of time to worsening or time to ST initiation will be used as ST initiation is considered representative of a negative outcome.

Table 1–2: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for MDS-UPDRS estimands

		MDS-UPDRS Part I-III sum score and Part I/II/III subscales	
Estimand	**Primary, key secondary and secondary**	**Key secondary**	**Secondary and supplementary**
Handling ICE (ST initiation)	**In the absence of ST initiation**	**ST initiation as part of the response**	**Regardless of ST initiation**
Population summary measure	**For Part I-III sum score**: Difference in slope of progression at 12 months (primary) and difference in mean at 12	**For Part III subscale:** Difference in RMET (time to confirmed *first* 5-point increase or ST	**For Part I-III sum score**: Difference in mean at 12 and 18 months (supplementary)

Table 1–2: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for MDS-UPDRS estimands

	and 18 months (key secondary and supplementary)[a] **For Part III subscale:** Difference in slope of progression at 12 months (key secondary) **For Part III subscale and Part III ePD subscore[a]:** Difference in mean at 12 and 18 months (key secondary) **For Part I/II subscale:** Difference in mean at 12 and 18 months (secondary)	initiation) within 18 months (key secondary) **For Part II subscale:** Difference in RMET (time to confirmed *first 3*-point increase or ST initiation) within 18 months (supplementary)	**For Part I/II/III subscales:** Difference in mean at 12 and 18 months (supplementary)
ICE & ICE-like protocol deviations			
ST initiation (Non-MAO-B Inhibitors)[c]	Hypothetical (set to missing & impute all post ICE visit data – MAR/ML)	Composite (ICE is part of the endpoint)	Treatment policy (Include post-ICE data)
ST initiation (MAO-B Inhibitor)[c]	Hypothetical (set to missing & impute all post ICE visit data – MAR/ML)	Composite (ICE is part of the endpoint)	Treatment policy (Include post-ICE data)
Treatment discontinuation (Lack of efficacy or AE)	Treatment policy (Include post-ICE data)	id	id
Treatment discontinuation (Other causes)	Treatment policy (Include post-ICE data)	id	id
Important ICE-like protocol deviations			
with long-term impact [d]	Hypothetical (set to missing & impute all post ICE visit data – MAR/ML)	Hypothetical (set to missing & treat as right-censored (non-event) at time of initial IPD)	id
With short-term impact [d]	Hypothetical (set to missing & impute single post-ICE visit data – MAR/ML)	Hypothetical (ignore)	id

Table 1–2: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for MDS-UPDRS estimands

Death or serious injury (all causes)	Actual missing: impute all post ICE visit data – MAR/ML	Actual missing: treat as right-censored (non-event) at time of termination	id
COVID-19-related ICEs without Treatment discontinuation or Study termination	Treatment policy (Include post-ICE data)	id	id
Missed visits, study termination and loss to follow-up			
Missed visit(s) (intermittent)[f]	Actual missing: impute visit data – MAR/ML	Actual missing: ignore	Actual missing: impute visit data – MAR/ML
Study termination: Treatment-related	Actual missing: impute all post termination visit data – MAR/ML[g]	Actual missing: Imputed as non-responder (event) to randomised treatment under MNAR[e]	Actual missing: impute all post termination visit data – MAR/ML
Study termination and loss to follow-up Other/unknown causes	Actual missing: impute all post termination visit data – MAR/ML	Actual missing: will be treated as right-censored (non-event) at the time of termination	Actual missing: impute all post termination visit data – MAR/ML

AE=Adverse Event; COVID-19 = Coronavirus Disease-2019; ICE=Intercurrent Event; id = idem, "the same" across rows; MAR = Missing at Random; ML = Maximum Likelihood; MDS-UPDRS=Movement Disorder Society-Unified Parkinson's Disease Rating Scale; RMET=Restricted Mean Event Time; ST = Symptomatic Treatment.

[a] For the estimands of MDS-UPDRS Part I-III sum score, Part III subscale and Part III ePD subscore, de-mediated difference in mean at Month12/18, the handling strategies outlined in this table for ST initiation and for study termination do not apply. Details on the handling strategies to be applied for ST initiation and study termination will be provided in Appendix 5.

[c] See Section 5.3.3.3 for how to determine which type of ST a medication is.

[d] The category that each protocol deviation falls into will be defined outside of the SAP, in the protocol deviation specifications or a separate document. Only important protocol deviations can be ICEs, but not all important protocol deviations will be defined as ICEs.

[e] Treatment-related study-termination is classed as an event in the RMET analysis for MDS-UPDRS Part III.

[f] This does not apply to missing baseline data. Baseline data will not be imputed.

[g] ST de-mediation approaches may implement alternative imputation strategies.

Note: In case of an event preventing a participant from attending a particular scheduled clinic visit with a MDS-UPDRS assessment (for example, due to COVID-19), this will be considered a minor protocol deviation and MDS-UPDRS data collected for this participant in clinic at future visits will be included in the analyses. Any imputation of the missing visit data will apply a MAR assumption.

Note: When MDS-UPDRS is collected at home rather than in clinic not all MDS-UPDRS data will be collected (rigidity assessments will be missing) and the Part III and total scores cannot be calculated. Any imputation of the missing visit data will apply a MAR assumption.

Note: Where missing data is to be imputed under a MAR assumption, whether in the case of either intermittent missing or of longitudinal missing (study termination or loss to follow-up), a ML approach will be used for all LMEM-based modelling.

Note: the Part II emerging symptoms estimand is not covered in this table. The ICE handling strategies for this estimand are covered in Section 5.3.2.1.7.

Table 1–3: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for other estimands

	DaT-SPECT mean Striatum SBR	ST intake	MoCA
Estimand	**Secondary**	**Secondary, other secondary and supplementary**	**Other secondary**
Handling ICE (ST initiation)	**Regardless of ST initiation**	**ST initiation/intake as the response**	**Regardless of ST initiation**
Population summary measure	Difference in mean change from Baseline at 12 months and 18 months (secondary)	Difference in RMET (time to ST initiation) within 18 months (secondary). Odds ratio of ST intake at 12 months (other secondary) and 18 months (secondary), Relative Risk of ST intake at 12 and 18 months (sensitivity), Hazard ratio of ST intake at 18 months (supplementary)	Difference in mean at 18 months (other secondary)
ICE & ICE-like protocol deviations			
ST initiation (MAO-B Inhibitors)[b]	Hypothetical (set to missing & impute all post ICE visit data – MAR/ML)	NA	Treatment policy (Include post-ICE data)
ST initiation (Non-MAO-B Inhibitors)[b]	Treatment policy (Include post-ICE data)	NA	Treatment policy (Include post-ICE data)
Treatment discontinuation (Lack of efficacy or AE)	Treatment policy (Include post-ICE data)	id	id
Treatment discontinuation (Other causes)	Treatment policy (Include post-ICE data)	id	id
Change in scanner	Treatment policy (Include post-ICE data)	NA	NA
Important ICE-like protocol deviations			

Table 1–3: Overview of handling strategies for ICEs, ICE-like protocol deviations and study termination or loss to follow-up for other estimands

With long-term impact [c]	Hypothetical (set to missing & impute all post ICE visit data – MAR/ML)	Hypothetical (set to missing & treat as right-censored (non-event) at time of initial IPD)	id
With short-term impact [c]	Hypothetical (set to missing & impute single post-ICE visit data – MAR/ML)	Hypothetical (ignore)	Treatment policy (Include post-ICE data)
Death or serious injury (all causes)	Hypothetical (set to missing & impute all post ICE visit data – MAR/ML)	Actual missing: treat as right-censored (non-event) at time of termination	id
COVID-19 related ICEs without Treatment discontinuation or study termination	Treatment policy (Include post-ICE data)	id	id
Missed visits, study termination and loss to follow-up			
Missed visit(s) (intermittent)[e]	Actual missing: impute visit data – MAR/ML	Actual missing: ignore	Actual missing: impute visit data – MAR/ML
Study termination: Treatment-related	Actual missing: impute all post termination visit data – MAR/ML	Actual missing: Imputed as non-response (event) to randomized treatment under MNAR[d]	Actual missing: impute all post termination visit data – MAR/ML
Study termination and loss to follow-up: Other/unknown causes	Actual missing: impute all post termination visit data – MAR/ML	Actual missing: will be treated as right-censored (non-event) at the time of termination	Actual missing: impute all post termination visit data – MAR/ML

AE=Adverse Event; COVID-19 = Coronavirus Disease-2019; DaT-SPECT=Dopamine Transporter Imaging with Single Photon Emission Computerized Tomography; ICE=Intercurrent Event; id = idem, "the same" across rows; MAR = Missing at Random; ML = Maximum Likelihood; MNAR = Missing Not at Random; MoCA=Montreal Cognitive Assessment; RMET=Restricted Mean Event Time; SBR=Specific Binding Ratio; ST=Symptomatic Treatment

[b]See Section 5.3.3.3 for how to determine which type of ST a medication is.

[c]The category that each protocol deviation falls into will be defined outside of the SAP, in the protocol deviation specifications or a separate document. Only important protocol deviations can be ICEs, but not all important protocol deviations will be defined as ICEs.

[d]Treatment-related study-termination is classed as an event (ie, non-response) in the RMET analysis for ST initiation.

[e]This does not apply to missing baseline data. Baseline data will not be imputed.

Note: Where missing data will be imputed under a MAR assumption, whether in the case of either intermittent missing or longitudinal missing (study termination or loss to follow-up), a ML approach will be used for all LMEM-based modelling.

1.2.1 Symptomatic treatment initiation

The main ICE affecting the interpretation of measurements will be initiation of ST.

MDS-UPDRS

Initiation of ST will be considered to impact the definition of the intervention effect of interest when measured using MDS-UPDRS Part I-III sum score as well as using MDS-UPDRS Part II and Part III subscales. Only Levodopa and Dopamine-agonist STs will be considered as having an impact on MDS-UPDRS, see Section 5.3.3.3 for the classification of ST types and Table 1–2 for how to handle other types of ST.

In the primary estimand analysis, the post-ICE data will be censored ('Hypothetical' strategy, assuming data is missing at random) and the data will be analyzed using a likelihood-based method of analysis, ie, a LMEM for longitudinal data/repeated measures with its implicit imputation approach (White, 2012) applied to all participants in the target population. This type of analysis will be referred to as 'in the absence of ST initiation'.

An alternative 'Hypothetical' handling strategy of this ICE will also be implemented as a supplementary estimand where data recorded under ST will be corrected using a de-mediation approach (see Section 5.3.2.1.4).

In a supplementary analysis to the primary estimand, the post-ICE data up to Month 18 will be included in the analysis ('Treatment policy' strategy) irrespective of whether the participant complied with the 12-hour ST washout: for all scheduled visits, study participants who start to receive ST over the study observation period will be asked to refrain from taking ST for at least 12 hours before the clinic visit and to bring the medication to the site. Visits where the participant did not comply with the 12-hour ST washout will be identified in the analysis datasets, but will not be censored from the analyses. If it is not possible to determine whether a participant completed the washout period for a particular visit, it will be assumed that they were compliant.

DaT-SPECT

A secondary endpoint will be the mean striatal dopamine receptor-specific binding ratio (SBR) as measured by DaT-SPECT. In this study, it will be assumed that DaT-SPECT signal is not affected by initiation or dose of levodopa (reviewed in Ikeda et al, 2019). Initiation of levodopa, dopamine agonists, and COMT inhibitors as well as change in dose or type will be considered NOT to impact the effect of interest for DaT-SPECT, and the post ICE data will be included in the main estimand analysis 'regardless of ST initiation' ("Treatment policy" strategy). Missing data due to scan not having taken place at the participant's or investigator's discretion will be treated as missing at random.

MOA-B inhibitors are not allowed in PD0053. However, if they are taken, a 'hypothetical' strategy will be used in all analyses for these medications and post-ICE data will be set to missing. Taking MAO-B is considered a protocol deviation which is also considered an ICE, and therefore also covered in Section 1.2.3.

1.2.2 Treatment discontinuation and treatment-related ICE-like protocol deviations

Participant-led or Investigator-led treatment discontinuation may be related to assigned investigational treatment (due to lack of efficacy / AEs) or unrelated.

Note that some of the participants who discontinue assigned investigational treatment due to lack of efficacy (whether on placebo or UCB0599) may want to start taking ST and therefore their data will be handled as such.

Treatment discontinuation will not be considered to impact the definition of the intervention effect of interest, and the post-ICE data will be included in the analyses ('Treatment policy' strategy).

Minor treatment-related protocol deviations such as IMP treatment non-compliance (missing a dose or taking a dose at a different time of the day) or minor drug administration error will not be considered to impact the definition of the intervention effect of interest, and the post-ICE data will be included in the analyses ('Treatment policy' strategy).

1.2.3 Other ICE-like protocol deviations

Some protocol deviations can be defined as ICEs. The protocol deviations specification document will clearly define the protocol deviations that are ICEs and the type of impact they will have.

For important protocol deviations considered to impact the effect of interest on the long-term, the deviation will be considered uninformative with respect to the treatment of interest and the post-ICE data will be removed from the analyses for all following visits and the post ICE missing data will be imputed ("Hypothetical" handling strategy) assuming (MAR) and using ML imputation.

For important protocol deviations considered to only impact the effect of interest in the short-term, e.g. a single visit, the deviation will be considered uninformative with respect to the treatment of interest and data for the impacted visit(s) will be removed from the impacted analyses and the post ICE missing data will be imputed ("Hypothetical" handling strategy) assuming (MAR) and using ML imputation.

Further details on protocol deviations are given in Section 5.1.2.2.

1.2.4 Death or serious injury

Local injuries eg, to arms/hands or legs/feet for MDS-UPDRS and to head for DaT-SPECT, or systemic acute conditions such as a stroke or an accident-related coma may prevent the taking of measurements. Treatment-related death or death due to other causes (including COVID-19) may occur during the study. Death due to PD is unlikely, as participants have been selected in the early stage of PD, although an accident may be the consequence of PD symptoms.

Both PD-related and non PD-related serious injury or death will be considered uninformative with respect to the definition of the effect of the intervention of interest, and the post-ICE data will be set to missing in the analyses ('Hypothetical' strategy). In practice, these events will be identified as serious adverse events that result in persistent or significant disability/incapacity or death.

1.2.5 Confirmed or suspected COVID-19

Confirmed or suspected cases of COVID-19 will not be considered to impact the definition of the intervention effect of interest, and the post-ICE data will be included in the analyses ('Treatment policy' strategy). This is assuming that the participant remains in the study.

This strategy only applies to confirmed or suspected COVID-19. Other ICEs or protocol deviations related to COVID-19 (treatment or study discontinuation) should be handled using the approaches outlined in the sections above.

Study participants are permitted to receive the COVID-19 vaccine at any point in the study. A treatment policy strategy will be used for this ICE where post-vaccination data is kept as part of the analyses.

1.2.6 Study termination and loss to follow-up

A participant may withdraw from the study at any time at his/her own request or may be withdrawn at any time at the discretion of the Investigator for safety, behavioral, IMP compliance, or administrative reasons. If a study participant withdraws or is withdrawn, he or she should be encouraged to perform the End-of-Treatment (EOT) Visit and the Safety Follow up Visit approximately 30 days after last dose of IMP.

Study termination will only be considered treatment-related if the reason for termination is 'lack of efficacy'. As a default approach, MAR will be assumed and the implicit imputation of likelihood-based linear models (for gaussian distributed outcomes), will be applied to any occurrence of study termination (including treatment-related termination) and the event will be ignored. For event-based outcomes, the approach will depend on whether the termination is considered to be treatment-related or not. For treatment-related termination, imputation assuming non-response to treatment will be applied (for example, for time to ST initiation we would impute treatment-related termination as a progression-related 'event'). For other causes of study termination (including loss to follow-up), data will be considered right censored at the time of termination (and ignored).

In well-designed clinical studies it is reasonable to assume that dropout patterns follow the MAR mechanism, although missing not at random (MNAR) data cannot be ruled out (Liu-Seifert, 2015). As an alternative approach, a different hypothetical strategy, such as a reference-based imputation (RBI) approach (for gaussian distributed outcomes) will be used when participants terminate the study for treatment-related reasons for MDS-UPDRS-based continuous endpoints estimands (Appendix 3).

1.3 Study design

PD0053 is a randomized, double-blind, placebo-controlled, 18-month Phase 2a study to evaluate the efficacy, safety, tolerability, and PK of orally administered UCB0599 in study participants with early PD who are not treated with symptomatic medications targeting motor symptoms of PD at the time of inclusion. Oral UCB0599 capsules or matching placebo capsules will be administered BID, approximately 12 hours apart.

The primary objective of the study is to demonstrate the superiority of UCB0599 over placebo with regard to clinical symptoms of disease progression over 12 and 18 months in this patient population. The difference between UCB0599 and placebo will be evaluated for both the low and high doses of UCB0599 (180mg/day and 360mg/day). In all analyses, the 2 dose levels will be assessed separately against placebo, there will be no pooling of the 2 dose levels, and they will not be compared to each other.

The comparison of the high dose of UCB0599 with placebo will be considered as the primary evaluation. If the comparison of the low dose and placebo gives a positive result on the primary estimand but the comparison of the high dose and placebo does not, these results will have to be interpreted with caution.

PD0053 includes a Screening Period of 3 to 6 weeks (including where available, a wearable sensor familiarization period for those participants who consent to its use), an 18-month Treatment Period (including, where available, a wearable sensor familiarization period for those participants who consent to its use after the Screening Period), and an SFU Period of 1 month. Study participants who complete the Treatment Period will have the option to transition into a dose-blinded extension study, PD0055. In this case, participants will not enter the SFU Period of PD0053.

The study will be conducted utilizing a partly decentralized model, ie, study visits may be composed of a combination of clinic visits and remote visits (except in France; for France-specific requirements, please refer to the protocol). During remote visits, study assessments will be conducted with the study participant from his/her home. This is offered in order to reduce study participant burden and encourage greater study participation. Further details on the decentralized model are given in the study protocol (Section 4.1.1).

However, all assessments necessary for the analyses of the Primary Estimand (i.e. MDS-UPDRS (Part I + III subscales) will be conducted on-site so that the same rater (on-site Clinician) can undertake the MDS-UPDRS assessments for given participant (at the relevant visits). Only safety, PK and exploratory efficacy assessments will be conducted during remote visits, see the schedule of assessments in protocol section 1.3 for details on the exact assessments to be performed.

Participants will be randomized to either UCB0599 360mg/day (180mg BID), UCB0599 180mg/day (90mg BID), or placebo when the IMP for low dose UCB0599 180mg/day (90mg BID) will be available to supply sites so that the final allocation ratio is 1:1:1. Approximately 645 participants will be screened to achieve 450 randomly assigned to study medication and 429 evaluable participants (accounting for study termination prior to month 3), for an estimated total of 143 evaluable participants per treatment group. To balance prognostic factors across treatment groups, randomization of study participants will be stratified using permutation blocks by gender.

The outcomes of this study will inform the design and dose selection for confirmatory studies and help to further validate novel endpoints and technologies and support strategic development program decisions.

The schedule of activities for this study can be found in protocol section 1.3 and the study schematic can be found in protocol section 1.2.

2 STATISTICAL HYPOTHESES

Since the main objective of this study is to provide proof of concept for the efficacy of UCB0599 in reducing disease progression in study participants with early-stage PD, the null hypothesis for all efficacy analyses is that there is no difference between UCB0599 high dose and placebo.

This is an exploratory study with only one primary estimand (efficacy), therefore no formal multiplicity adjustments are planned. P-values from any statistical models will be presented but they should not be over-interpreted; any decision-making and interpretation should take the context of the analysis, the effect size and the confidence intervals into account (Wasserstein, 2016).

3 SAMPLE SIZE DETERMINATION

The sample size for this study was determined based on the primary efficacy estimand as detailed in the sections below. Approximately 645 participants will be screened to achieve 450 participants randomly assigned to study medication and 429 evaluable participants, for an estimated total of 143 evaluable participants per treatment group. Participants who terminate the study (withdraw consent) or are lost to follow-up will not be replaced.

The primary comparison of interest is UCB0599 high dose (360mg/day) vs placebo. Given the limitations in our understanding of the PK/PD properties of the compound, the sample size calculations presented below apply to both comparisons (UCB0599 high dose [360mg/day] and placebo and UCB0599 low dose [180mg/day] and placebo) as similar assumptions for both doses are made with respect to variability and impact on progression.

3.1 Primary efficacy estimand

The primary efficacy objective of PD0053 is to show superiority of UCB0599 over placebo with regards to clinical symptoms of disease progression over 12 months in participants diagnosed with early-stage PD. To measure clinical symptoms of disease progression, the MDS-UPDRS Part I-III sum score was selected as the variable of interest for the primary estimand and the population-level summary was chosen as the difference in the population mean slope of progression between UCB0599 and Placebo up to 12 months or up to the time of ST initiation, whichever comes first.

Even though the duration of the study is 18 months, the sample size was estimated to ensure 90% power at 12 months, where the proportion of participants who have initiated ST is still limited. Beyond 12 months, the number of participants under ST is expected to increase with significant potential for bias between the UCB0599 and placebo groups (should the compound be active), and power decrease due to data loss. The primary analysis of MDS-UPDRS is therefore planned at 12 months, and additional analyses will be performed at 18 months including post-ST initiation data in the analysis.

In order to derive the parameters of interest for sample size estimation, an Estimator similar to the one planned for the primary estimand (a linear mixed effects model (LMEM) for longitudinal data/repeated measures, with time as a continuous fixed effect, ie, assuming linear development over 12 months, and identical random effects to account for within-participant correlation between repeated observations – see Section 5.3.1) was applied to a subset of the Parkinson Progression Markers Initiative (PPMI) 1.0 de novo cohort, an observational study of early-stage PD participants (Parkinson's Progression Markers Initiative, 2018). The model was not adjusted for Baseline covariates.

Only patients enrolled in the early PD cohort from which participants with scan without evidence for dopaminergic deficit (SWEDD) were excluded, who were aged between 40 and 80 and who had been diagnosed with the disease for at most two years at screening were included in this analysis. This age range is larger than the one chosen for the study inclusion criteria (40 to 75 years), but since the variability is likely to be larger for those aged between 75 and 80, this approach is conservative. In addition, the following patients were also excluded:

(1) patients for whom MDS-UPDRS at Baseline was not available;

(2) patients who were already on ST at 3 months;

(3) patients who had their MDS-UPDRS measured twice or less over the first 12 months of the study.

This left 361 out of the 423 early-stage PD participants in the dataset (participants with SWEDD excluded). This participant sample was assumed to be representative of the POC study placebo arm up to 12 months.

For the purpose of sample size estimation, a closed-form sample size formula for longitudinal data was used (Ahn, 2014 – Chapter 5, formula 5.19).

Number of participants per arm:

$$\frac{\left\{\left(1+\frac{1}{\lambda}\right)*\left\{(1-\rho)+\left(r_\tau * N_1 * Var(T)\right)\right\}*Z^2_{\left(1-\frac{\alpha}{2};\beta\right)}\right\}}{N_1 * Var(T) * \Delta^2}$$

Where λ is the allocation ratio, ρ is the intra-class correlation, r_τ is the ratio of the random slope variance component to the sum of the other variance terms (the random intercept variance component σ_u^2 and the residual variance σ_e^2), ie, $r_\tau = \sigma_{ut}^2 / \sigma^2$ with $\sigma^2 = \sigma_u^2 + \sigma_e^2$. N_1 is the number of timepoints, Var(T) is the population variance of the timepoints and Δ is the absolute effect size standardized to σ (ie, $\Delta = (0.3 * \beta_t)/\sigma$, where β_t is the time coefficient from the LMEM applied to the PPMI data).

The causes for correlation can be modelled using random effects. The random effects will then induce the correlation structure of the marginal model. The formula used here assumes that the correlation between the participant-specific random intercept and the participant-specific random effect of time (random slope) is null and that $\sigma_{ut}^2 \approx 0$, resulting in a model equivalent to a random intercept model where intra-class correlation $(\rho) \approx \sigma_u^2 / \sigma^2$ for all participants and timepoints. For the purpose of sample size calculation, the correlation between the random effects was set to 0 when running the LMEM. In practice, this correlation was estimated to be 0.43.

Therefore, in this case the resulting unconditional variance-covariance matrix resembles a compound symmetry working correlation matrix, ie, equivalent to a random intercept model with the variance of observations equaling σ^2 and covariance between observations equaling σ_u^2.

In addition, the formula assumes the variance components will not differ by treatment group.

The frequency of MDS-UPDRS assessments was assigned to be bi-monthly. The population variance of the time points, Var(T), was weighted for MDS-UPDRS data loss over time assuming an exponential loss function and a monotonic missing pattern.

All analyses were performed in R version 3.6.1 (2019-07-05, The R Foundation for Statistical Computing) using the lme4 package.

Based on the PPMI data, the parameters were estimated to be:

Table 3–1: Parameter estimates based on PPMI data

σ_u^2	154.0
σ_{ut}^2	0.43
σ_e^2	31.1
β_t	0.90
λ	1
$\boldsymbol{Z^2_{(1-\frac{\alpha}{2};\beta)}}$	2.56^2
N_1	7
N_1*Var(T)	85.57[a]

[a] for 35% data loss at 12 months

3.2 Incorporating data loss

In the PPMI dataset, 10.7% of patients had a missing value recorded at their (initial) 3-month MDS-UPDRS assessment (42/391 patients aged 40-80 years), but only 3% (11/391) were truly lost to follow-up, see details below. Here, data loss cannot be explained by study termination related to study medication as no investigational treatment was applied. Another 2.6% of patients had initiated ST by 3 months (10/391).

A 5% data loss due to study termination by 3 months (all causes) will be assumed in this study, based on expert knowledge. The sample size will be inflated accordingly (see details below). Loss to follow-up and loss to ST initiation by 3 months will be considered minimal in the study setting and will be ignored.

At 12 months, 58% (228/391) of the patients in the PPMI dataset had initiated ST (Figure 6–1 in Appendix 3, Section 6.3 illustrates this), and a least one missing value had been recorded at a single timepoint in 82% (320/391) of the patients. However, in the PPMI dataset missing data does not follow a monotonic missing pattern.

For example, out of the 32 patients who had a missing value recorded at 3 months, only 34% (11/32) had missing data for all 3 follow-up assessments up to 12 months, and 15% (5/32) had missing data on 2 follow-up assessments, while 41% (13/32) had only 1 further missing record, and 3 patients (10%) had valid records at 6, 9 and 12 months. Similarly, among the 27% (107/391) of patients who had a missing value recorded for the first time at 6 months, only 5% (5/107) of them had missing data for both 9 and 12 months follow-up, while among the 7% (27/391) of patients who had a missing value recorded for the first time at 9 months, 74% (20/27) of them had missing data at the 12-month follow-up. In addition, 9% (36/391) of patients had a first missing value recorded at 12 months, with 17% (6/36) of these having a missing value at 18 months (the presence of missing data was not investigated beyond 18 months for the purpose herewith).

If we assume that only those PPMI patients with a monotonic pattern of missing data would be representative of the missing data situation observed in a RCT setting, we would estimate another 10.7% (11+5+20+6 = 42/391) data loss at 12 months unrelated to initiation of ST (Figure 6–1 in Appendix 3, Section 6.3 illustrates this).

Based on these observations, 35% data loss was assumed in the study at 12 months:

- 20% due to participants initiating ST, as they will be in a controlled setting and encouraged to delay ST initiation for as long as it is deemed medically acceptable;
- 5% due to participants lost to follow-up as well as to death and serious injury;
- 10% due to participants terminating the study (all causes).

Data loss was modelled using an exponential loss function assuming no data loss at 3 months (Figure 6–2 in appendix 3, Section 6.3 illustrates this).

Based on the parameters obtained from the analysis performed on the restricted PPMI dataset for the placebo arm, a sample size of 143 participants per arm (N = 429 total), randomized in a 1:1:1 ratio to UCB0599 high dose, low dose or placebo, would provide 90% power to detect a minimum of expected 30% decrease in population mean slope of PD disease progression (equivalent to 0.27 MDS-UPDRS points/month or 3.2 points/year) in the active treatment groups compared with placebo over 12 months of treatment, with 10% type I error (1-sided). In reference to MDS-UPDRS, the smallest improvement considered clinically relevant has been shown to be -6.7 points mean change from baseline (95% CI -9.4; -3.9, Makkos et al, 2018).

The sample size will be inflated by 5% (N = 21) to account for participants terminating the study by 3 months, leading to a final sample size of 450 participants.

Since it is assumed that between 12 and 18 months another 10% of data (at a minimum) will be lost due to ST initiation and another 10% due to study termination and loss to follow-up, analyses at 18 months will be carried out irrespective of intake of ST (all data included).

3.3 Power for key secondary efficacy estimand (MDS-UPDRS Part III)

A key secondary estimand to the primary efficacy objective of PD0053 is the difference in population mean slope of progression between UCB0599 and placebo at 12 months in the absence of ST initiation for the MDS-UPDRS Part III subscale.

Using the same approach as described in Section 3.1, the sample size estimated to achieve 90% power on the primary efficacy estimand (MDS-UPDRS Part I-III sum score) ie, 143 participants per arm, would provide 90% power to detect a minimum of expected 35% relative decrease in population mean slope of progression on the MDS-UPDRS Part III subscale (equivalent to 0.19 points/month or 2.3 points/year) in the UCB0599 treatment groups compared with placebo over 12 months of treatment, with 10% type I error (1-sided).

For the MDS-UPDRS Part III subscale, a decrease of 3.25 points mean change from baseline (95%CI -4.32; -2.17) has been identified as the minimal clinically important improvement (Horvath et al., 2015).

3.4 Power for secondary efficacy estimand (DaT-SPECT mean striatum SBR)

A secondary efficacy objective of PD0053 is to show superiority of UCB0599 over placebo with regards to neurodegeneration over 18 months in participants diagnosed with early-stage PD.

To assess neurodegeneration, the DaT-SPECT mean Striatum SBR was selected as a participant-level variable and the population-level summary chosen as the difference between UCB0599 and placebo in the population mean change from Baseline (screening) at 18 months.

Change from Baseline (screening) data was derived from the PPMI 1.0 de novo cohort DaT-SPECT mean Striatum SBR data. The same PPMI patients were selected as for MDS-UPDRS. In addition, patients for whom DaT-SPECT data was not available either at screening, 12 or 24 months were excluded from the analysis to derive the parameters of interest for sample size estimation. This left 297 out of the 423 early PD participants in the PPMI dataset.

As DaT-SPECT data was not available at 18 months in PPMI, we used the midpoint between the change from Baseline at 12 and 24-months as an estimate of the change from Baseline to be observed at 18 months. With respect to variability, we used the estimate of standard deviation for the 24-month data.

Using a two-sample t-test based sample size formula and the parameters for the placebo arm obtained from the PPMI 1.0 de novo DaT-SPECT data analysis, the sample size estimated to achieve 90% power on the primary efficacy estimand (MDS-UPDRS Part I-III sum score) ie, 143 participants per arm, would provide 90% power to detect a minimum of expected 37.5% relative increase in population mean change from Baseline in DaT-SPECT signal (equivalent to 0.067 SBR unit) in the UCB0599 groups vs placebo group at 18 months, with 10% type I error (1-sided), and assuming 25% data loss (ie, an extra 10% loss to occur between 12 and 18 months due to study termination and loss to follow-up at 18 months with effective N = 107/arm). Data loss due to ST initiation does not apply for the assessment of the DaT-SPECT signal as the assumption is being made that levodopa and dopamine agonists do not affect the neurodegeneration of dopamine neurons in the midbrain.

To increase information regarding DaT-SPECT SBR signal, participants intending to terminate the study (assumed to be 10% by 12 months) should have a scan before study termination (as part of the EOT visit). Participants who choose to withdraw from the study at or after 6 months and accept to have a scan at time of study termination will have their scan analysed as if it was taken at 12 months.

Residual data attrition due to study termination and loss to follow-up may be imbalanced between treatment arms by the time participants reach the 18-month timepoint, creating selection bias. This imbalance could happen in different directions depending on the balance of efficacy to AEs in the study. Compared to the placebo arm:

- a larger number of participants could be lost in the UCB0599 arms compared to placebo due to treatment-related AEs;
- a smaller number of participants could be lost in the UCB0599 arms compared to placebo due to treatment-related efficacy.

Consequently, different numbers and subtypes of participants per arm may continue to the end of the study. We estimate we would be left with an imbalance between arms of 5% due to data loss to follow-up, and another 5% data loss due to study termination before 6 months, as well as another 5% in the period between 12 and 18 months.

For these participants, we will control for bias at the analysis stage using implicit imputation of likelihood-based approaches (see Section 5.3).

4 POPULATIONS FOR ANALYSIS

The following analysis sets will be used:

- *All Study Participants Set (ASPS):* All study participants who sign the ICF. Treatment assignment in the ASPS will be based on treatment received.
- *Randomized Set (RS):* All study participants who are randomized. Treatment assignment in the RS will be based on treatment received.
- *Safety Set (SS):* All randomized study participants who receive at least a partial dose of study medication. Treatment assignment in the SS will be based on treatment received.
- *Full Analysis Set (FAS):* All randomized study participants who receive at least a partial dose of study medication, have at least 1 post-Baseline assessment. This is *any non-missing* post-baseline assessment, including unscheduled assessments. The FAS will be used for all efficacy analyses, and analyses will be conducted based on randomized treatment. This analysis set will exclude participants who do not meet key inclusion/exclusion criteria (see appendix 6), these will include criteria related to: age and disease duration at time of informed consent/baseline, Hoehn and Yahr stage at screening, evidence of dopamine transporter deficit (from screening DaT-SPECT) and ST status at screening.
- *Pharmacokinetic Set (PKS):* All study participants in the Safety Set who have at least 1 observable PK measurement. The PKS will be used for PK analyses and treatment assignment in this analysis set will be based on actual treatment received.

The FAS will be used for all analyses with ICEs and other protocol deviations to be handled according to the approach specified for each estimand (see Sections 1.2). For a given endpoint, the analysis models will exclude any participant with a missing assessment at baseline for that particular endpoint.

5 STATISTICAL ANALYSES

5.1 General considerations

All tables, figures and listings (TFLs), including statistical evaluation, will be produced by Parexel using SAS version 9.4 or higher (SAS Institute, Cary, North Carolina, USA) unless stated otherwise. The analyses covered in Section 5.3.1.4.1 and Section 5.3.2.1.4 will be produced by UCB. Analysis data will adhere to Clinical Data Interchange Standards Consortium (CDISC) guidance documents for Analysis Data Model (ADaM) and follow the UCB interpretation.

TFLs will be presented by treatment group and visit as applicable. Listings of all documented and calculated data will be presented by treatment group, study participant and visit. Data collected at unscheduled visits will be included in listings but not in any summary tables unless explicitly stated otherwise.

Continuous variables

Unless stated otherwise, summary statistics will be presented for continuous variables including number of participants (n), mean, standard deviation (SD), median, minimum and maximum. When summarizing efficacy endpoints, the inter quartile range (IQR) will also be included. Geometric mean, geometric coefficient of variation (CV) and 95% CI for the geometric mean will also be presented in the summaries of UCB5099 and N-oxide metabolite concentration data. Further details on presenting concentration data will be given in Section 5.5.1.

Generally, means (arithmetic or geometric), standard deviations, medians and IQRs will use one decimal place more than the original data and confidence intervals will use 1 decimal place more than the value which they are constructed around. Depending on the data format this could be significant figures rather than decimal places. Minimum and maximum will use the same format as the original data. In listings, if the data is coming directly from the CRF it should be presented exactly as captured. For derived variables, data should be reported using 1 decimal place more than the values which they are calculated from.

For all continuous efficacy endpoints, summary statistics will be produced for both observed results and changes from Baseline.

Time to event variables

For time to event variables, descriptive summary tables will include the number and percentage of study participants who reach the endpoint by each visit.

Categorical variables

Categorical variables will be summarized using frequency counts and percentages. Unless otherwise stated, the denominator for the percentage calculations will be based on the number of study participants in the respective analysis set, treatment group and visit (as applicable) with non-missing data.

Unless otherwise noted, all percentages will be displayed to 1 decimal place. No percentage will be displayed for zero counts, and no decimal will be presented when the percentage is 100%.

Estimands

Estimands will be described as a function of their attributes: (1) the population of interest, (2) the treatment effect of interest, (3) the participant-level variable (or endpoint) of interest, (4) the specification of how ICEs are reflected in the scientific question of interest, and (5) the population-level summary for the variable/endpoint. (See Table 1–1, Table 1–2 and Table 1–3 for the estimands and ICE handling strategies to be used in this study).

5.1.1 Date imputation

Partial dates may be imputed for the following reasons:

- Classification of adverse events (AEs) as treatment-emergent
- Classification of medications as prior or concomitant
- Calculation of duration of exposure
- Calculation of duration of AEs
- Calculation of Parkinson's disease duration
- Determining start/end date of EST

Imputed dates will not be shown in the listings; all dates will be displayed as reported in the database.

The following rules will be applied for partial AE and concomitant medication start dates when classifying as treatment-emergent or prior/concomitant and also for determining AE duration:

- If only the month and year are specified and the month and year of the first dose of IMP is not the same as the month and year of the start date then the 1st of the month will be used, or the date of Screening visit 1 if this is later (if the latter imputation results in an end date that is earlier than the start date, then the 1st of the month will be used).
- If only the month and year are specified and the month and year of the first dose of IMP is the same as the month and year of the start date, then the date of the first dose of IMP will be used. If this results in an imputed start date that is after the specified end date, then the 1st of the month will be used, or the date of Screening visit 1 if this is later (if the latter imputation results in an end date that is earlier than the start date, then the 1st of the month will be used).

- If only the year is specified, and the year of the first dose of IMP is not the same as the year of the start date then January 01 will be used.
- If only the year is specified, and the year of the first dose of IMP is the same as the year of the start date, then the date of the first dose of IMP will be used. If this results in an imputed start date that is after the specified end date, then January 01, or the date of Screening visit 1 if this is later will be used (if the latter imputation results in an end date that is earlier than the start date, then January 01 will be used). If the imputed date is the date of first dose of IMP then the event will be regarded as treatment-emergent or the medication as concomitant.
- If the start date is completely unknown, the start date will not be imputed.

The following rules will be applied to partial AE and concomitant medication stop dates when classifying as treatment-emergent or prior/concomitant:

- If only the month and year are specified, then the last day of the month will be used.
- If only the year is specified, then December 31 of the known year will be used.
- If the stop date is completely unknown, the stop date will not be imputed.

If the above imputation still leads to a missing AE date:

Table 5–1: Calculation rules for duration of adverse events

Data availability	Onset date/time	Outcome date/time	Calculation rules
Complete data	D1	D2	Duration = D2 – D1 + 1
Start date missing	--	D2	Duration = D2 – D0 + 1 For a study participant in the SS, D0 is the date of first administration of IMP and for study participants who were randomized and not dosed, or for screen failures, D0 is the date of Screening Visit 1.
End date missing	D1	--	If the stop date is missing, duration will not be calculated.

Rules for calculating duration of exposure (UCB0599 or placebo)

For partial study medication stop dates, the last day of the month will be used if only day is missing. If the start date is missing or more than just the day is missing for the medication stop date, exposure duration will not be calculated. If partial dates are recorded for date of first and last dose of study medication this should be queried prior to applying any imputation rules.

Rules for calculating duration of Parkinson's Disease

Partial date of diagnosis dates will be imputed with the 15th of the month if only day is missing, if both day and month are missing the date will be imputed with 30th June (ie, the middle of the year).

Rules for determining start/end date of ST

Partial ST initiation dates will be imputed with the 15th of the month (if only day is missing) for the purposes of analyzing ST data (determining when ST started for ICE handling, and for calculating LEDD). This imputation rule will also apply to partial ST end dates. If month and/or year is missing, no imputation will be performed. If partial ST dates are recorded, these should be queried prior to applying any imputation rules.

5.1.2 General study level definitions

5.1.2.1 Analysis time points

5.1.2.1.1 Relative day

The relative day of an event will be derived with the date of first dose of study medication as reference.

Relative day for an event or measurement occurring before the date of first dose in the study will be calculated as follows:

$$Relative\ Day = Event\ Date - Date\ of\ First\ Dose$$

The relative day for an event or measurement occurring on the date of first dose is day 1. The relative day for an event or measurement occurring on or after the first dose and before the date of the last dose will be calculated as follows:

$$Relative\ Day = (Event\ Date - Date\ of\ first\ Dose) + 1$$

For events or measurements occurring after the date of last dose of study medication (within the SFU period), relative day will be prefixed with "+" in the data listings and will be calculated as follows:

$$Relative\ Day = +(Event\ Date - Date\ of\ Last\ Dose)$$

There is no relative Day 0. Relative day will not be calculated for partial dates. In such cases, relative day should be presented as "--" in the relevant data listing.

5.1.2.1.2 End date of the treatment period

The end date of the treatment period will be either the date of the month 18 EOT visit (visit 15) for participants who complete the treatment period, or the date of the EOT visit for participants who discontinue the study during the treatment period (same procedures as visit 15 for completers). If a study participant does not have an EOT visit then the date of the last scheduled or unscheduled visit during the treatment period or the date of last known dose of study medication (whichever is later) will define the end date of the treatment period.

5.1.2.1.3 Study periods

The total duration of the study per participant is approximately 21 months. The end of the study is defined as the date of the last visit of the last participant in the study. The following study periods are defined for this study:

Screening period:

The screening period will last 3 to 6 weeks and includes 2 visits to the investigational site (Visit 1 and Visit 2). At Visit 1 the participants provide informed consent and undergo screening to assess eligibility for the study. At Visit 2, additional assessments to verify eligibility, including DaT-SPECT imaging, will be performed. Participants consenting to its use will also receive a wearable sensor (the wearable sensor) for 2-3 weeks for familiarization at home. This is to address potential technical issues and to collect stable Baseline data for the sensor. The sensor will be finally returned by the participant to the clinic on Visit 15 (M18 visit, or Early Termination Visit in case of early study termination).

Treatment period:

The treatment period will last for 18 months and may include both investigational site visits and telemedicine video calls with the site (see schedule of activities in protocol section 1.3 for more detail). During video calls a research nurse will support the participant at the participant's home. The treatment period ends with Visit 15 (EOT).

During the treatment period, participants will receive either UCB0599 (360mg/day), UCB0599 (180mg/day) or placebo with a 1:1:1 randomization ratio.

Safety follow up period:

The safety follow up will last for approximately 30 days and includes 1 visit (the Safety Follow-Up (SFU) visit). The SFU visit may be a telehealth video call with the investigator, supported by the research nurse. For those participants who do not roll over to the extension study (PD0055), the study ends with the SFU Visit approximately 30 days after the last dose. Participants who enter the extension study will not enter the SFU period and their final visit will be visit 15.

Early termination:

In case of early termination of a participant's treatment, the participant will be asked to attend the End of Treatment (EOT) and the SFU visit (30 days after the last dose) and will not be eligible for the dose-blinded extension study (PD0055).

5.1.2.1.4 Mapping of assessments performed after early termination

Study participants who discontinue the study for any reason will be asked to attend the End of Treatment (EOT) and the SFU Visit (30 days after the last dose) and will not be eligible for the extension study (PD0055). Participants will be encouraged to attend these two visits as soon as possible after last dose of study drug.

The following rules will apply regarding the inclusion of data obtained at the EOT visit for participants who discontinue:

- If the early EOT visit occurs at the same time as the next scheduled visit, the results will be included with all other participants' results from that visit in summary tables
- If the above is not true, the results will be mapped to the nearest scheduled visit following the last scheduled visit where assessments were performed (use the earliest visit if equidistant)
- If results are then mapped to a visit where the assessment is not performed, these results will be listed only and not included in any summaries
- Mapped results will be included in any statistical analyses (where applicable)

These rules apply to all data, with the exception of DaT-SPECT scan results. As scans are only planned three times throughout the study, the above mapping rules will not be used as we would often map to visits where a scan is not scheduled. For rules for mapping EOT visit DaT-SPECT scans for participants who discontinue early, and also for out of window DaT-SPECT scans, see Section 5.3.2.3.

The results from these early end of treatment visits will be displayed as the mapped visit and flagged in the data listings. Early SFU visit results will not be mapped to another visit but will be flagged in data listings. Results from early SFU visits will be listed but will not be included in summary tables and figures, unless explicitly stated otherwise (for example, impact of COVID-19 summaries).

For all statistical analyses, data for participants who terminate early will be handled using the pre-specified ICE strategies (see Section 1.2 for further details). If the data is to be included in the analysis, the mapping rules above will be used.

5.1.2.1.5 Definition of baseline values

In general, Baseline values will be determined from the latest non-missing value collected prior to the first dose of study medication, unless otherwise stated for a specific analysis in later sections.

If day 1 pre-dose data is collected per the schedule of activities then this will be used as the Baseline value, if not the latest screening data will be used. All day 1 assessments will be assumed to have been taken pre-dose. If day 1 pre-dose data is supposed to be collected but is not for any reason, this should be investigated before automatically using the latest screening value.

5.1.2.2 Protocol deviations

Important protocol deviations (IPDs) are deviations from the protocol which could potentially have a meaningful impact on study conduct or on either the primary or key secondary outcome(s) for an individual study participant. The criteria for identifying such protocol deviations will be defined within the IPD specifications document.

Important protocol deviations will be categorized as follows:

- Inclusion/exclusion criteria deviations
- Incorrect treatment or dose administered
- Procedural non-compliance
- Prohibited concomitant medication use
- Withdrawal criteria deviation

All protocol deviations will be reviewed as part of the ongoing data cleaning process and all important deviations will be identified and documented before unblinding. Important protocol deviations will be listed and summarized. Three data evaluation meetings (DEMs) will be held throughout this study, with the final DEM occurring just before database lock. The purpose of these DEM reviews will be to review all protocol deviations and check the quality of the data. The reviews will also help decide how to manage problems in the study participants' data (eg, missing values and withdrawals).

Accepted deviations from scheduled time points will be described in the appropriate documents and included in the eTMF. After the pre-analysis review, resolution of all issues, and documentation of all decisions (including inclusion into each of the analysis sets) at the final DEM, the database will be locked.

Some protocol deviations will be pre-defined as ICEs and the handling of these deviations is discussed in Section 1.2.

5.1.2.3 Treatment assignment and treatment groups

Participants will be randomized in a 1:1:1 ratio to receive either UCB0599 high dose (360mg/day), UCB0599 low dose (180mg/day) or placebo. All listings and summaries will be presented by treatment group unless stated otherwise. Appendix 4 outlines which summaries should present an overall summary as well as data summarized by treatment group.

Table 5–2: Treatment group descriptions

Full Description	Data Display Label
Placebo	Placebo
UCB0599 360mg/day (180mg BID)	Minzasolmin 360mg/day
UCB0599 180mg/day (90mg BID)	Minzasolmin 180mg/day
UCB0599 pooled high dose (360mg/day) and low dose (180mg/day)	Minzasolmin Total

If after unblinding it is determined that participants received treatment other than what they were randomized to, then for baseline characteristics, safety and PK analyses participants will be allocated to the treatment they predominantly received (ASPS, RS, SS and PKS assignment will be based on actual treatment). Treatment assignment for the FAS will be allocated based on randomized treatment and not actual treatment received, all efficacy analyses will be based on the FAS. A sensitivity analysis will be performed for the primary estimand where actual treatment assignment is used.

5.1.2.4 Center pooling strategy

The data from different centers will be pooled for all analyses. Data for some endpoints will be summarized by region/country as described in Section 5.1.2.6.

5.1.2.5 Coding dictionaries

Adverse events and medical history will be coded using the latest available version of the Medical Dictionary for Regulatory Activities (MedDRA®). Medications will be coded according to the latest available version of the World Health Organization Drug Dictionary (WHODrug-Global B3 format). Medical procedures will not be coded.

5.1.2.6 Multicenter studies

Since treatment assignment will not necessarily be balanced across countries or sites, no statistical analyses will be carried out to investigate center effects. However, for the primary endpoint and for some pre-specified secondary endpoints, summary statistics will be produced by region (Europe and North America).

For each region summary statistics for both observed results and changes from Baseline by visit (Baseline, months 12 and 18 where applicable) and treatment group will be produced for the following endpoints:

- MDS-UPDRS Part I-III sum score;
- MDS-UPDRS Part III;
- MDS-UPDRS Part III ePD subscore (UCB in-house)
- MDS-UPDRS Part III – First 5-point Increase;
- MDS-UPDRS Part II;
- MDS-UPDRS Part II – First 3-point Increase;
- MDS-UPDRS Part I;
- MoCA;
- DaT-SPECT mean striatum SBR;
- Levodopa cumulative daily dose (summary of changes from Baseline not needed for this endpoint).

These summaries will also present the number of participants included in each treatment group by region.

In addition, for each country and region, the following counts will be summarized by visit and treatment group, overall and by gender:

- The number of participants on ST.

In addition, in response to the COVID-19 pandemic, the number of participants who contracted COVID-19 during the study (further details given in Section 5.2) will be summarized by treatment group and visit, overall and by country/region.

5.2 Participant dispositions

The number of study participants who started the study, completed the study, prematurely discontinued, and who enter the extension study will be presented by treatment group and overall. The reasons for discontinuation will also be summarized. This summary of disposition and discontinuation will be based on the RS. A study participant who completed the study is defined as a study participant who has a completed status in the study termination CRF. Since the primary estimand analysis in this study will be done using the first 12 months of data for each participant, the number of completers and the number of premature discontinuations will be summarized separately up to this time (in the same summary table). Participants who discontinue early are considered to have completed month 12 if their study termination date falls in or after the month 12 visit window.

The number and percentage of study participants who discontinue due to AEs will be summarized by treatment group, based on the RS. This summary will be used for European Union Drug Regulating Authorities Clinical Trials (EudraCT) reporting. This table will also present (separately) discontinuations due to AEs prior to the month 12 visit.

The number and percentage of study participants included in each of the analysis sets will be summarized by treatment group. Numbers will be presented for all analysis sets (including the ASPS) but percentages will be calculated based on the RS for the purpose of this summary. Additionally, a listing based on the SS will be produced for participants excluded from the FAS, including the reason for exclusion.

Screen failure reasons will be summarized for the ASPS. A separate summary will also be produced summarizing the reasons for ineligibility for those who screen fail due to being ineligible. A listing of study participants who did not meet study eligibility criteria will also be presented for this analysis set.

In addition, the following listings will be presented:

- Disposition of study participants (ASPS)
- Study discontinuation (RS)
- Visit dates (RS)
- Study participant analysis sets (ASPS)
- Rescreened Participants (ASPS)

The listing of disposition of study participants will include the date of informed consent, date of first and last dose of IMP, date of premature termination and primary reason (if applicable), date of final contact and whether the participant entered the extension study.

The listing of study discontinuation will include the reason for discontinuation and the number of days on IMP.

The number of days on IMP will be calculated as follows:

$$Number\ of\ Days\ on\ IMP = (Date\ of\ Last\ Dose\ Received - Date\ of\ First\ Dose\ Received) + 1$$

Additionally, a summary of the impact of COVID-19 (for any reason) will be produced by country and treatment group for the RS. This information will also be listed for the ASPS. Potential impacts that will be included in these summaries are missed visits or doses, visits performed out of window or through a different modality, and temporary/permanent study or drug discontinuation. Relationship to COVID-19 will also be included in these outputs, detailing whether the impact is due to confirmed infection, suspected infection or general COVID-19 related circumstances.

Two additional outputs will be produced for use in the Plain Language Summary (and are not required for the CSR):

- Summary table of disposition of participants screened (showing the dates of the first participant in and out at each site, as well as analysis set counts by site)
- Listing of actual treatment assignment by site and participant

5.3 Efficacy estimands and endpoints

All efficacy data will be listed based on the FAS. All summaries and analyses for efficacy estimands and endpoints will be presented for the FAS.

Statistical outputs

For all statistical models fitted to the efficacy data, diagnostic plots will be included in the statistical output documents. The statistical outputs for all models will include the following treatment group comparisons:

a. UCB0599 high dose (360mg/day) versus placebo,
b. UCB0599 low dose (180mg/day) versus placebo,

For some key analyses (defined elsewhere), the model output summary table will also present the percentage change in mean slope or in mean difference at Month 12/18 for each dose-level against placebo (ie, the difference in slope/mean divided by the placebo slope/mean multiplied by 100).

For some key analyses (defined elsewhere), the mean slope or the mean difference at Month 12/18 for each dose-level against placebo will be depicted in forest plots with either corresponding 95% CIs or 80% CIs, the latter to match the type I error used in the trial sample size estimation.

MDS-UPDRS score derivation

For all MDS-UPDRS analyses, MDS-UPDRS Part I-III and subscales (Part I, Part II, Part III) sum scores will only be calculated if responses are available for all questionnaire items. If one item response is missing, for example in Part I, Part II and Part III sum scores can still be calculated but Part I and Part I-III sum scores will not be calculable. However, missing responses to individual items within the questionnaire are not expected.

To calculate the sum scores, the response scores to the following items will be summed:

- Part I: items 1.1 to 1.13, sum score ranges from 0 to 52;
- Part II: items 2.1 to 2.13, sum score ranges from 0 to 52;
- Part III: items 3.1 to 3.18 (for Part III some items will be tested on both side of the body and on the upper as well as the lower limb – all responses will be summed to get the Part III total score, 33 items in total), sum score ranges from 0 to 132;
- Part I-III: items 1.1 to 1.13, 2.1 to 2.13 and 3.1 to 3.18, sum score ranges from 0 to 236.

When MDS-UPDRS is collected at home rather than in clinic not all MDS-UPDRS data will be collected (rigidity assessments will be missing) and the Part III and Part I-III sum scores cannot/will not be calculated. Part I and Part II sum scores can still be calculated and used in any analyses, assuming all item responses are collected.

Symptomatic Treatment Status and other ICEs

For all summary tables and listings where ST status is used or presented, the assumption is made that once a participant starts ST, they remain on ST for all study visits after their initiation date.

For summary tables that are being presented for participants who are not yet on ST, this means not yet on ST by a particular visit. This means that the number of participants contributing to the summary will differ by visit in the table, as more participants may have initiated ST by the following visit.

For plots that are displayed by ST initiation timing, the following categories will be defined for participants who started ST:

- By month 2 (inclusive);
- between month 2 and 4 (inclusive);
- between month 4 and 6 (inclusive);
- between month 6 and 8 (inclusive);
- between month 8 and 10 (inclusive);
- between month 10 and 12 (inclusive);
- between month 12 and 18 (inclusive);
- not started by 18 months/EOT.

To support the efficacy analyses presented in this section, a (cumulative) summary table presenting the number and percentage of participants with each ICE (see Section 1.2 for a full list of the ICEs defined in this study) will be produced by visit.

Modelling considerations

Baseline covariates

For any analyses adjusting for continuous covariates, these covariates will be mean centered to aid interpretability (the intercept term can be interpreted as the expected value of the response when the dependent variables are set to their means). When mean centering covariates, the calculated mean will be based on the data used for that estimand analysis. All efficacy models will include age at baseline as a covariate. Since this study was randomized stratifying for gender, gender should be included as a covariate in all efficacy models unless explicitly stated otherwise.

As stated in Section 4, for all analyses, a participant should only be included in the model if they have a baseline score for the endpoint being analyzed.

Estimation considerations and Convergence issues

If convergence problems are encountered, model (any) will be run without adjusting for age at baseline. If convergence issues persist, check that all variance components are positive or not too close to zero. If a variance component is negative or close to zero, then attempt to simplify the structure of the relevant variance-covariance matrix. If convergence issues are solved by using a simplified variance-covariance structure, age at baseline may be added back into the model if appropriate.

For Mixed Effects Models, the following steps will be implemented:

- Linear random coefficients model with Time as continuous: the most common issue will be for the estimate of covariance between the subject and the subject-time random coefficients. In this case, set the corresponding covariance in the G variance covariance matrix (RANDOM statement) to 0 to obtain a separate variance component for each random effect (TYPE = VC) and re-run the model:

$$G = \begin{pmatrix} \sigma_{participant} & 0 \\ 0 & \sigma_{participant-time} \end{pmatrix}$$

- ANCOVA/Repeated measures model with Time as categorical: set the structure of the variance-covariance matrix to a compound symmetry structure (instead of ARH1 in the case of MDS-UPDRS or unstructured in the case of DaT-SPECT).

For details, refer to Brown and Prescott, 2014.

5.3.1 Primary estimand analysis: MDS-UPDRS Part I-III sum score slope of progression over 12 months, in the absence of ST initiation

For the primary estimand analysis, only data from the first 12 months of the study will be considered. The primary estimand is for the primary efficacy objective, which is to demonstrate the superiority of UCB0599 over placebo with regards to clinical symptoms of disease progression over 12 months.

5.3.1.1 Definition of endpoint

The primary endpoint for each participant is slope of progression in MDS-UPDRS Part I-III sum score measured bi-monthly over 12 months or up to the time of ST initiation, including Baseline score and scores at 2, 4, 6, 8, 10 and 12 months.

Descriptive statistics

Although the primary estimand is based only on data up to month 12, descriptive statistics discussed here for MDS-UPDRS Part I-III sum score will be presented for the entire 18-month period.

Summary tables presenting the observed mean and mean change from Baseline in MDS-UPDRS Part I-III sum score by treatment group and visit will be produced. This summary will be produced overall and by ST status, where ST status is defined as having or not having initiated ST.

Plots of individual MDS-UPDRS Part I-III sum score over time by treatment group and gender will be presented, color-coded to clearly indicate when a participant is and is not on ST. These plots will be grouped by ST initiation timing category.

In addition to these individual plots, plots of observed mean MDS-UPDRS Part I-III sum score and mean change from baseline will be produced by treatment group and ST initiation timing category, overall and by gender. These plots will include error bars (+/- SD) and all treatment groups will be overlaid on the same plot.

All planned summary tables and plots of MDS-UPDRS Part I-III sum score will also be produced for the sum scores of each individual subscale (Part I, II and III).

5.3.1.2 Main analytical approach

Details on the ICE handling for this estimand can be found in Section 1.2. All data recorded after a participant initiates ST will be censored.

A Linear Mixed Effects Model (LMEM) here, a random coefficients model for longitudinal data will be applied to the data up to the time of ST initiation. MDS-UPDRS Part I-III sum score will be the dependent variable, Baseline MDS-UPDRS data will be part of the dependent variable rather than a covariate in the model. Time since baseline in months, as a continuous variable and the treatment by time interaction (fixed slope) will be fitted as fixed effects (Twisk et al, 2018), gender and age at baseline will be included in the model as covariates. To adjust for the correlation between repeated observations from each participant, a participant-specific random intercept and a participant-specific random slope (ie, participant by time interaction or random time effect) will be fitted as random effects (Twisk, 2013). The random effects will induce the correlation structure in the marginal model (ie, the model not conditional on random effects following integrating them out).

A fixed effect for treatment will not be included in this model. This implicitly assumes that the Baseline values for all three groups are equal and reflected in the intercept of the model. The treatment effects can therefore be directly obtained from the treatment by time interactions (slope or rate of progression).

The population-level summary for this estimand will be the difference in population mean slope of progression in MDS-UPDRS Part I-III sum score over 12 months in the absence of ST initiation between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo.

The estimates of these will come from the treatment by time interaction from the above model. These estimates are to be interpreted as the unit increase in MDS-UPDRS Part I-III sum score per month, after adjusting for gender and age at baseline. A summary table presenting these estimates with the corresponding 95% confidence intervals will be produced. A repeat of this summary table will also be produced presenting the model estimates to 1 decimal place only, for use in the study Plain Language Summary (this table is not required for the CSR).

A plot displaying the adjusted mean slope and 95% confidence interval for each treatment group from this model will be produced.

5.3.1.3 Sensitivity analyses

The primary analysis may be repeated using additional prognostic covariates (e.g., Parkinsonian age/digital twin). The results will be part of a separate report and may not be part of the CSR.

5.3.1.4 Supplementary approaches

5.3.1.4.1 Supplementary analysis in the absence of ST initiation: de-mediated difference in mean at 12 months

Missing data will be imputed using multiple imputation (MI) according to randomized treatment.

A 'hypothetical' handling strategy of ST initiation will be implemented, where all data (up to Month 12) recorded after a participant initiates ST will be corrected using a de-mediation approach.

For further details on the MI and ST de-mediation approaches, see appendix 5.

The population-level summary of interest will be the ST de-mediated difference in target population mean MDS-UPDRS Part I-III sum score at Month 12 as if the participants had not initiated ST between the following treatment arms:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all ST de-mediated data recorded up to Month 12. Treatment, time (visit, categorical) and treatment by time (interaction term between treatment and visit) will be fitted as fixed effects, gender, age at Baseline and Baseline MDS-UPDRS Part I-III will be fitted as covariates. A heterogenous auto-regressive (ARH1) variance-covariance matrix will be fitted to account for the repeated measures within subject.

A summary table presenting the estimated ST de-mediated treatment effects of interest, i.e. difference in means at Month 12 (based on the treatment by time interactions) and corresponding 95% CIs will be produced.

A plot displaying the estimated ST de-mediated means and 95% confidence interval for each treatment group and each time point up to Month 12 will be produced.

5.3.1.4.2 Supplementary analyses (regardless of ST initiation)

As a supplementary analysis, MDS-UPDRS Part I-III sum score will be analyzed using a different ICE handling strategy with respect to ST. Here a "treatment policy" strategy will be used, where post-ST data will be included in the analysis.

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all data recorded up to Month 12. This LMEM will use the same modelling approach described in Section 5.3.1.4.1.

A summary table presenting the estimated treatment effects (based on the treatment by time interactions) and corresponding 95% CIs will be produced, with the estimates at 8, 10 and 12 months as the main results to be interpreted. A plot displaying the adjusted mean and 95% confidence interval for each treatment group from this model will also be produced.

5.3.1.4.3 Supplementary analysis in the absence of ST initiation: de-mediated difference in slope over 12 months

A LMEM (random coefficient model) for longitudinal data (model described in section 5.3.1.2) will also be applied to the MDS-UPDRS Part I-III sum score ST de-mediated data (ST de-mediation described briefly in section 5.3.1.4.1 with further details in appendix 5) to estimate the difference in ST de-mediated mean slopes between treatment groups.

A summary table presenting these estimates with the corresponding 95% confidence intervals will be produced. A plot displaying the adjusted mean slope and 95% confidence interval for each treatment group from this model will be produced.

5.3.2 Secondary estimands/endpoint analyses

The secondary estimands will be split into 4 sections based on the study objectives: clinical motor symptoms (MDS-UPDRS-related), clinical non-motor symptoms (MoCA), neurodegeneration (DaT-SPECT and total alpha-syn) and initiation/intake of ST. Some secondary estimands are defined over 12 months and others are defined over 18 months. It will be clearly stated which time points should be used in any analyses throughout this section.

5.3.2.1 Secondary estimands: clinical symptoms (MDS-UPDRS)

These secondary estimands fall under the primary objective of the study, which is to demonstrate the superiority of UCB0599 over placebo with regards to clinical symptoms of disease progression over both 12 and 18 months.

As secondary estimands MDS-UPDRS Part I-III sum score will be analyzed over 18 months and the MDS-UPDRS subscales will be analyzed individually.

In addition, a sub-score based on the MDS-UPDRS Part III and targeted at the early-stage PD population will also be analysed.

For all MDS-UPDRS analyses the treatment comparisons of interest are:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo.

5.3.2.1.1 MDS-UPDRS Part I-III sum score at 18 months, in the absence of ST initiation

Definitions of endpoint

The endpoint for each participant is MDS-UPDRS Part I-III sum score measured bi-monthly over 18 months or up to the time of ST initiation, including Baseline score and scores at 2, 4, 6, 8, 10, 12, 14, 16 and 18 months.

Main analytical approach

Details on the ICE handling for this estimand can be found in Section 1.2. All data recorded after a participant initiates ST will be censored.

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all data recorded up to Month 18. This LMEM will use the same modelling approach described in Section 5.3.1.4.1.

A summary table presenting the treatment effect estimates (based on the treatment by time interactions) with the corresponding 95% confidence intervals will be produced. The estimates at 18 months are the main results to be interpreted.

A plot displaying the adjusted mean and 95% confidence interval for each treatment group from this model will be produced.

Supplementary analyses (regardless of ST initiation)

As a supplementary analysis, the model above will be fitted again using a different ICE handling strategy for ST initiation. Here the participant-level endpoint will be MDS-UPDRS Part I-III sum score measured bi-monthly over 18 months, regardless of ST initiation. Here ST will be handled using a "treatment policy" strategy where post-ICE data is kept as part of the analysis. The same summary outputs will be produced as described for the main analysis.

5.3.2.1.2 MDS-UPDRS Part III subscale: slope of progression over 12 months

Since the Part III subscale captures the majority of the disability scale measured by the MDS-UPDRS sum score in early PD, the same analyses as planned for the primary estimand will be carried out for this subscale at 12 months.

Definition of endpoint

The primary endpoint for each participant is slope of progression in MDS-UPDRS Part III measured bi-monthly over 12 months or up to the time of ST initiation, including Baseline score and scores at 2, 4, 6, 8, 10 and 12 months.

Main analytical approach

This endpoint will be analyzed in the exact same way as the primary estimand, using the same ICE handling strategy (post-ST initiation data will be censored) and the same model (see Section 5.3.1.2).

The treatment effect estimates will come from the treatment by time interaction. These estimates are to be interpreted as the unit increase in MDS-UPDRS III per month, after adjusting for gender. A summary table presenting these estimates with the corresponding 95% confidence intervals will be produced.

A plot displaying the adjusted mean and 95% confidence interval for each treatment group from this model will also be produced.

Sensitivity analysis

The primary analysis may be repeated using additional prognostic covariates (e.g., Parkinsonian age/digital twin). The results will be part of a separate report and may not be part of the CSR.

Supplementary analyses (regardless of ST initiation)

As a supplementary analysis, MDS-UPDRS Part III will be analyzed using a different ICE handling strategy with respect to ST initiation. Here a "treatment policy" strategy will be used, where post-ST initiation data will be included in the analysis.

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all data recorded up to Month 12. This LMEM will use the same modelling approach described in Section 5.3.1.4.1 (baseline Part III score will be used in place of Part I-III sum score).

A summary table presenting the estimated treatment effect (based on the treatment by time interactions) and corresponding 95% CIs will be produced, with the estimates at 8, 10 and 12 months as the main results to be interpreted. A plot displaying the adjusted mean and 95% confidence interval for each treatment group from this model will also be produced.

5.3.2.1.3 MDS-UPDRS Part III subscale: difference in mean at 12 and at 18 months

Definition of endpoint

The endpoint for each participant is MDS-UPDRS Part III score measured bi-monthly over 18 months or up to the start of ST, including Baseline score and scores at 2, 4, 6, 8, 10, 12, 14, 16 and 18 months.

Main analytical approach

Details on the ICE handling for this estimands can be found in Section 1.2. All data recorded after a participant initiates ST will be censored.

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all data recorded up to Month 18. This LMEM will use the same modelling approach described in Section 5.3.1.4.1 (baseline Part III score will be used in place of Part I-III sum score).

A summary table presenting the treatment effect estimates (based on the treatment by time interaction) with the corresponding 95% confidence intervals will be produced. The estimates at Month 12 and Month 18 are the main results to be interpreted.

A plot displaying the adjusted mean and 95% confidence interval for each treatment group from this model will be produced.

Supplementary analyses (regardless of ST initiation)

As a supplementary analysis, MDS-UPDRS Part III will be analyzed using a different ICE handling strategy with respect to ST initiation. Here, a "treatment policy" strategy will be used, where post-ST initiation data will be included in the analysis. The same modelling approach as described above for the main analysis will be used.

5.3.2.1.4 MDS-UPDRS Part III subscale and ePD subscore: de-mediated difference in mean at 12 and at 18 months

Definition of endpoints

MDS-UPDRS Part III subscale and Part III ePD subscore measured bimonthly over 18 months, including Baseline score, and scores at 2, 4, 6, 8, 10, 12, 14, 16, and 18 months.

The analysis below will be conducted using all data recorded up to Month 18, for the two endpoints (subscale / ePD subscore).

Main analytical Approach

Missing data will be imputed using multiple imputation (MI) according to randomized treatment.

A 'hypothetical' handling strategy of ST initiation will be implemented, where all data (up to Month 18) recorded after a participant initiates ST will be corrected using a de-mediation approach.

For further details on the MI and ST de-mediation approaches, see appendix 5.

The population-level summary of interest will be the ST de-mediated difference in target population mean MDS-UPDRS Part III subscale/ePD subscore at Month 12 and Month 18 as if the participants had not initiated ST between the following treatment arms:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all ST de-mediated data recorded up to Month 18. Treatment, time (visit, categorical) and treatment by time (interaction term between treatment and visit) will be fitted as fixed effects, gender, age at Baseline and Baseline MDS-UPDRS Part III subscale (ePD subscore) will be fitted as covariates. A heterogenous auto-regressive (ARH1) variance-covariance matrix will be fitted to account for the repeated measures within subject.

A summary table presenting the ST de-mediated estimates for the treatment effects of interest, i.e. difference in means at Month 12 / Month 18 (based on the treatment by time interactions) and corresponding 95% confidence intervals will be presented.

A plot displaying the ST de-mediated estimated means and 95% confidence interval for each treatment group and each time point up to Month 18 will be produced.

Supplementary analyses (in the absence of ST initiation – ST de-mediated)

A LMEM (random coefficient model) for longitudinal data (model described in section 5.3.1.2) will also be applied to the MDS-UPDRS Part III subscale (ePD subscore) ST de-mediated data (ST de-mediation described briefly in section 5.3.1.4.1 with further details in section 6.5) to estimate the difference in ST de-mediated mean slopes between treatment groups.

A summary table presenting these estimates with the corresponding 95% confidence intervals will be produced. A plot displaying the adjusted mean slope and 95% confidence interval for each treatment group from this model will be produced.

5.3.2.1.5 MDS-UPDRS Part III subscale: time to worsening within 18 months

Definitions of endpoints

The participant-level endpoints will be the time from Baseline to the participant's *first* 5-point increase in MDS-UPDRS Part III subscale (or to last observed time prior to ST initiation) within the 18-month period, where the 5-point increase is confirmed at the next visit record, or ST initiated before the next visit record.

Main analytical approach

Full details on the ICE handling for this estimand can be found in Section 1.2. The main ICE will be initiation of ST, this will be handled using a "composite" strategy where initiation of ST is part of the endpoint definition. Study termination and loss to follow up related to treatment "lack of efficacy" will be handled by using non-responder imputation and be considered an event as part of the composite handling strategy. All other occurrences of study termination and loss to follow up will be right-censored, i.e., analysis only takes account of events which occurred within the observed duration.

The population level summary will be the difference in target population restricted mean event time (RMET) up to the last observed event (or last observed time prior to ST initiation) between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo

In this study, the RMET can be interpreted as the average time to the *first* 5-point increase in MDS-UPDRS Part III subscale within the study period or to initiation of ST, whichever comes first. This summary was chosen as it has been shown to be robust when the proportional hazards assumption does not hold, it can be estimated even under heavy censoring, and has an easy clinical interpretation (Royston and Parmar, 2013).

The RMET can be estimated as the area under a Kaplan-Meier (KM) survival curve up to a pre-specified point in time (Royston & Parmar, 2013) or in other words, the integral of the survival

function over the study period. However, a modelling approach will need to be used here so that we can adjust for gender and age at baseline. As the main analysis for time to the *first* 5-point increase in MDS-UPDRS Part III subscale, RMET from Baseline to the last observed event time will be estimated. This will be done by fitting a generalized linear model with a linear link function for the RMET, with gender and age at baseline as covariates and treatment group as the effect of interest. This model will be fitted using the inverse probability censoring weighting (IPCW) method of estimation, this method uses Kaplan-Meier estimation to obtain weights and it has been shown that weighting in this way provides an unbiased estimate for an adjusted survival curve (Calkins, 2018). By default, this approach assumes that the right-censoring mechanism is homogeneous among all participants, which may not be a valid assumption in this study; censoring in the context of this analysis will come from both ST initiation and study termination and it is possible that both may be higher or lower in the UCB0599 arms compared to placebo. This is especially true for ST initiation, if UCB0599 shows signs of efficacy it is likely that a lower number of participants will begin taking ST in these treatment groups. To counter this potential bias, we will obtain treatment group-specific weights by applying the KM estimation method separately to each group (ie, using treatment group as a stratification variable in the model).

A summary table presenting the adjusted difference in RMET of the *first* 5-point increase in MDS-UPDRS Part III subscale between each UCB0599 treatment group and placebo and the corresponding 95% CIs will be produced, this summary will also include the model estimates of adjusted RMET for each treatment group.

To support the interpretation of these analyses, a summary table will be produced showing the number and percentage of study participants who reach this endpoint by treatment group and visit. A listing will also be produced showing date and type of event (ie, 4 options: confirmed first 5-point increase, first 5-point increase followed by ST initiation, ST initiation only, or treatment-related "lack of efficacy" study termination) for each participant, based on the FAS.

Sensitivity analyses

The RMET will be estimated for each combination of treatment group and gender as the area under the Kaplan-Meier "survival" curve up to the last observed event time (ie, without using a modelling approach). A plot of these Kaplan-Meier curves will be produced, with the RMET estimates also included in the figure. These results can be used to assess the proportional hazards assumption as well as for comparison to the results using the modelling approach. Results from the log rank test (testing the null hypothesis that there is no difference between the six survival curves) will also be displayed on this plot.

5.3.2.1.6 MDS-UPDRS Part II subscale: difference in mean at 12 and 18 months

Definitions of endpoints

The endpoint for each participant is MDS-UPDRS Part II score measured bi-monthly over 18 months or up to the start of ST, including Baseline score and scores at 2, 4, 6, 8, 10, 12, 14, 16 and 18 months.

Main analytical approach

Details on the ICE handling for these estimands can be found in Section 1.2. All data recorded after a participant initiates ST will be censored.

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all data recorded up to Month 18. This LMEM will use the same modelling approach described in Section 5.3.1.4.1 (baseline Part II score will be used in place of Part I-III sum score).

A summary table presenting treatment effect estimates (based on the treatment by time interactions) and corresponding 95% CIs will be produced. The estimates at 8, 10, 12 and 18 months will be the main results to be interpreted.

Plots displaying the adjusted means and 95% confidence intervals for each treatment group from the above model will be produced.

Sensitivity analyses

This model will be refit using log-transformed data, as the normality assumption for the residuals does not appear to hold for this subscale when analyzing the PPMI data (UCB internal observation). If the participants Part II sum score is 0, the log-transformed score will be calculated as log(0.1). The same summary outputs will be produced as for the main analysis, results will be back-transformed to aid interpretability.

Supplementary analyses (regardless of ST initiation)

As a supplementary analysis, the model will be fitted again using a different ICE handling strategy for ST initiation. Here ST will be handled using a "treatment policy" strategy where post-ICE data is kept as part of the analysis. This supplementary analysis will be carried out on log-transformed data. The same summary outputs will be produced as described for the main analysis.

5.3.2.1.7 MDS-UPDRS Part II subscale: emerging symptoms within 18 months

Emerging symptoms

Definitions of endpoint

The emerging symptoms based on MDS-UPDRS Part II (Tosin et al., 2022) for individual participants are derived using the following steps:

1) Identify the symptoms not present at Baseline, ie, MDS-UPDRS Part II items with score '0' at Baseline.
2) For the items identified in step 1, calculate change from Baseline at each visit. The participant is considered to have an emerging symptom for the item (or event), if the change from Baseline for the item is greater than 0 for 2 consecutive visits or if symptomatic therapy was initiated between the first visit identified as showing a change greater than 0 and the following visit with a recorded score. The magnitude of change from Baseline will not be considered to determine the emerging symptom.

3) The participant-level endpoint will be derived as the total number of emerging symptoms identified in step 2.
4) The participant-level annualized rate of events is calculated as total number of emerging symptoms calculated in step 2 divided by total duration of follow-up.

Main analytical approach

The main ICE will be initiation of ST, this will be handled using a "composite" strategy. When the ST is initiated before the occurrence of the event this will be ignored. Initiation of ST between the emergence of a symptom for an item and the following visit is part of the endpoint definition.

All occurrences of important protocol deviations with long-term impact, death, study termination and loss to follow up, will be ignored (ie, rate of occurrence of the events is calculated only for the observed duration therefore introducing a small bias on the estimate of annualized rate of events). Missed visits and important protocol deviations with short-term impact will also be ignored. Treatment discontinuation and COVID-19 related ICEs will be handled in the same way as described for other MDS-UPDRS endpoints.

The population-level summary will be the ratio of annualized rate of emerging symptoms based on MDS-UPDRS Part II over 18 months period between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo

The annualized rates will be estimated using a negative binomial regression model with a fixed treatment effect and the log of the observed duration in study as off-set. The model will include gender and age as covariates and log of observed duration as off-set.

Estimates for the treatment effects of interest (see above) over the 18-month period adjusted for gender and age at Baseline with corresponding 95% CI.

To support the interpretation of these analyses, a summary table will be produced showing the number and percentage of study participants who experienced 0, 1, 2, … and up to 13 emerging symptoms/events over 18 month by treatment group.

Supplementary Analyses:

Alternative ST initiation handling strategy

The same analysis will be repeated as above but using a "Treatment policy" handling strategy for ST initiation, i.e. the definition of the endpoint will be altered so that ST initiation will no longer be considered when determining an emerging symptom.

Alternative Part II endpoint (symptoms worsening)

Definitions of endpoints

The participant-level endpoints will be the time from Baseline to the participant's *first* 3-point increase in MDS-UPDRS Part II subscale (or to last observed time prior to ST initiation) within the 18-month period, where the 3-point increase is confirmed at the next visit record or ST is initiated before the next visit record.

ICE Handling Strategies

Full details on the ICEs handling for this estimand can be found in Section 1.2. The main ICE will be initiation of ST, this will be handled using a "composite" strategy where initiation of ST is part of the endpoint definition. Study termination and loss to follow up related to treatment "lack of efficacy" will be handled by using non-responder imputation (considered as an event). All other occurrences of study termination and loss to follow up will be right-censored, i.e., analysis only takes account of events which occurred within the observed duration.

Population Summary

The population level summary will be the difference in target population restricted mean event time (RMET) up to the last observed event (or last observed time prior to ST initiation) between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo

In this study, the RMET can be interpreted as the average time to the above defined event-based endpoint.

Estimator

This analysis will use the same modelling approach as described for the MDS-UPDRS Part III subscale time to worsening analysis (see Section 5.3.2.1.5 for further details).

Estimates

A summary table presenting the adjusted difference in RMET of the *first 3*-point increase in MDS-UPDRS Part II subscale between each UCB0599 treatment group and placebo and the corresponding 95% CIs will be produced, this summary will also include the model estimates of adjusted RMET for each treatment group.

To support the interpretation of these analyses, a summary table will be produced showing the number and percentage of study participants who reach this endpoint by treatment group and visit. A listing will also be produced showing date and type of event (i.e. 4 options: confirmed *first* 3-point increase, *first* 3-point increase followed by ST initiation, ST initiation only, or treatment-related "lack of efficacy" study termination) for each participant, based on the FAS.

Sensitivity analyses

The RMET will be estimated for each combination of treatment group and gender as the area under the Kaplan-Meier "survival" curve up to the last observed event time (ie, without using a modelling approach). A plot of these Kaplan-Meier curves will be produced, with the RMET estimates also included in the figure. These results can be used to assess the proportional hazards assumption as well as for comparison to the results using the modelling approach. Results from the log rank test (testing the null hypothesis that there is no difference between the six survival curves) will also be displayed on this plot.

5.3.2.1.8 MDS-UPDRS Part I subscale: difference in mean at 12 and 18 months

Definitions of endpoints

The endpoint for each participant is MDS-UPDRS Part I score measured bi-monthly over 18 months or up to the start of ST, including Baseline score and scores at 2, 4, 6, 8, 10, 12, 14, 16 and 18 months.

Main analytical approach

Details on the ICE handling for these estimands can be found in Section 1.2. All data recorded after a participant initiates ST will be censored.

A LMEM for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to all data recorded up to Month 18. This LMEM will use the same modelling approach described in Section 5.3.1.4.1 (baseline Part I score will be used in place of Part I-III sum score).

A summary table presenting treatment effect estimates (based on the treatment by time interactions) and corresponding 95% CIs will be produced. For the analysis over 12 months, the estimates at 8, 10 and 12 months will be the main results to be interpreted. The estimates at 8, 10, 12 and 18 months will be the main results to be interpreted.

Plots displaying the adjusted means and 95% confidence intervals for each treatment group from this model will be produced.

Sensitivity analyses

As a sensitivity analysis, the model will be refit using log-transformed data, as the normality assumption for the residuals does not appear to hold for this subscale when analyzing the PPMI data (UCB internal observation). If the participants Part I sum score is 0, the log-transformed score will be calculated as log(0.1). The same summary outputs will be produced as for the main analysis.

Other secondary estimand analysis (regardless of ST initiation)

For this secondary estimand, the model will be fitted again using a different ICE handling strategy for ST initiation. Here ST will be handled using a "treatment policy" strategy where post-ICE data is kept as part of the analysis. This analysis will be carried out on log-transformed data. The same summary outputs will be produced as described for the main analysis.

5.3.2.2 Secondary estimands: clinical symptoms (MoCA)

These secondary estimands fall under the primary objective of the study, which is to demonstrate the superiority of UCB0599 over placebo with regards to clinical symptoms of disease progression over 18 months.

The MoCA (Nasreddine et al, 2005) assesses different cognitive domains (visuospatial/executive, naming, memory, attention, language, abstraction, delayed recall, and orientation). Participants are assessed on a 30-point scale. If 'years of schooling' is missing or unknown on the MoCA questionnaire, it will be assumed that the participant has had ≤ 12 years of schooling. A score of 26 or above is considered normal. For all listings, summaries and analyses of MoCA data the derived total score will be used rather than the total score collected in the database.

5.3.2.2.1 Definition of endpoints

The endpoint for each participant will be observed MoCA score at 18 months, regardless of ST initiation.

Descriptive statistics for observed and change from Baseline scores at 18 months will be presented. Summary tables will be produced by treatment group and visit overall and by ST status, where ST status is defined as having or not having initiated ST.

Plots of individual trajectories over time will also be produced. To make these plots interpretable with our large sample size, plots will be produced by treatment group, gender and age category (using the categories defined in Section 5.3.3.1). Trajectories will be color coded to clearly indicate when a participant initiates ST.

5.3.2.2.2 Main analytical approach

Here ST initiation will be handled using a "Treatment policy" approach, where post-ST initiation data is kept as part of the analysis. This assumes that the symptoms measured by the MoCA questionnaire are not affected by the recommended ST. Further details on the ICE handling for this estimand can be found in Section 1.2.

An analysis of covariance (ANCOVA) will be applied to all observed 18-months data ie, completer analysis using SAS PROC MIXED. Treatment will be included in the model as a categorical fixed effect, gender, age at Baseline and Baseline MoCA data will be fitted as covariates. The population-level summary of interest will be the difference in target population mean observed MoCA score at 18 months regardless of ST initiation between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo.

A summary table presenting the model estimates of these differences and corresponding 95% confidence intervals will be presented.

5.3.2.3 Secondary estimands: neurodegeneration

These secondary estimands are for the secondary objective of demonstrating superiority of UCB0599 over placebo with regards to neurodegeneration of dopaminergic neurons over both 12 and 18 months.

Participants who do not have a Month 12 DaT-SPECT scan will only have their Baseline and 18 months' data included in analyses over 18 months. Therefore, they will not contribute to the estimates of treatment difference at Month 12. For participants who do not have a DaT-SPECT scan taken at baseline, their historical scan results can be used as baseline in these analyses.

The following rules will be followed for early termination and out of window scans for summarizing and analyzing this data:

- For scans that occur after a participants screening/historical scan but prior to a participants Month 6 visit, results from these scans will be listed only;
- Scans that occur between Month 6 and Month 14 (inclusive) will be mapped to Month 12;
- Scans that occur after Month 14 will be mapped to Month 18.

5.3.2.3.1 Definition of endpoints

The whole striatum will be calculated as the average of the specific binding ratio (SBR) data values for the four following "small" regions: left caudate small, left putamen small, right caudate small and right putamen small. The SBR will be calculated for each region with the occipital cortex as a reference region, where a lower SBR indicates worse disease. The following formula will be used to calculate this:

(Average (Small region) – Average (Occipital region))/ (Average (Occipital region))

At both Month 12 and Month 18, the participant-level endpoint will be change from Baseline in DaT-SPECT whole striatum SBR at that time, regardless of ST initiation. For DaT-SPECT, Baseline is the value recorded at Screening visit 2 (or data from a historical DaT-SPECT scan acquired within 3 months of Screening Visit 1).

Descriptive statistics

Descriptive statistics for observed results and changes from Baseline in DaT-SPECT mean Striatum SBR at both Month 12 and Month 18 will be presented. Summary tables will be produced by treatment group and visit (Month 12 and Month 18), overall and by ST status, where ST status is defined as having or not having initiated ST.

Plots of individual trajectories over time will also be produced. To make these plots interpretable with our large sample size, plots will be produced by treatment group, gender and age category (using the categories defined in Section 5.3.3.1). Trajectories will be color coded to clearly indicate when a participant initiates ST.

In addition to these individual plots, longitudinal plots of mean observed DaT-SPECT mean Striatum SBR and mean change from baseline will be produced by treatment group, overall and by gender. These plots will include error bars (+/- SD) and all treatment groups will be overlaid on the same plot.

5.3.2.3.2 Main analytical approach

Details on the ICE handling for this estimand can be found in Section 1.2. It will be assumed that DaT-SPECT signal is not affected by ST initiation.

A LMEM for longitudinal analysis of covariance on change from Baseline will be applied to all data recorded up to Month 18, i.e. on 3 mean Striatum SBR values for each participant: screening, month 12 and month 18. Treatment, time (visit, categorical) and treatment by time (interaction term between treatment and visit) will be fitted as fixed effects, gender, age at Baseline and Baseline DaT-SPECT mean Striatum SBR will be fitted as covariates. Participant will be included in the model as a random effect.

The population-level summary of interest will be the difference in target population mean change from Baseline in DaT-SPECT mean Striatum SBR at Month 12 and Month 18 regardless of ST initiation between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo.

These estimates will come from both the treatment effects and the treatment by time interactions from the model; in this analysis the treatment terms alone reflect the treatment effect at 12 months. A summary table presenting the estimates of the treatment effects at Month 12 and Month 18 and corresponding 95% confidence intervals from this model will be presented. Since there is only one random effect in this model, the intra-class correlation coefficient (ICC) can be calculated and interpreted as the proportion of the total variance that is accounted for by variation within participants.

A plot displaying the adjusted mean and 95% confidence interval for each treatment group from this model (LMEM over 18 months) will be produced.

5.3.2.3.3 Sensitivity Analyses

- The model described in the previous section will be run on the original data (as opposed to the change from baseline data).
- An addition LMEM will be fitted to the original data, with Baseline data included as part of the response/dependent variable, rather than as a covariate, and time as continuous (instead of categorical) so as to estimate the slope between baseline and Month 18).

For these sensitivity analyses, the same summary tables & figures will be produced as described for the main analyses.

5.3.2.4 Secondary estimands: ST Initiation/Intake

A secondary efficacy objective of PD0053 is to assess the effect of UCB0599 vs placebo with regard to intake of ST over 18 months. This will be assessed by looking at time to initiation of ST and at the number of participants taking ST at 12 and 18 months.

See protocol Section 6.5.3 for details on which STs are permitted in this study.

To support these analyses, a summary table will be produced showing the number and percentage of study participants who initiate ST by treatment group and visit, overall and by gender. A listing and summary of concomitant ST medications, following the same format as standard concomitant medication outputs, will also be produced.

5.3.2.4.1 Time to ST Initiation

Definitions of endpoints

The participant-level endpoint will be time from Baseline to start of ST or last censored time observation within the 18-month period.

Main analytical approach

Full details on the ICE handling for this estimand can be found in Section 1.2. The main ICEs for this estimand will be study termination and loss to follow up (see details in Table 1–3).

The population level summary will be the difference in target population restricted mean event time (RMET) up to the last observed event time within the 18-month period between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo.

In this study, the RMET can be interpreted as the average time until ST initiation within the study period. This summary was chosen as it has been shown to be robust when the proportional hazards assumption does not hold, can be estimated even under heavy censoring, and has an easy clinical interpretation (Royston and Parmar, 2013).

As the main analysis for time to ST initiation, RMET from Baseline to the last observed event time will be estimated. This analysis will use the same modelling approach as described for the MDS-UPDRS Part III subscale time to worsening analysis (see Section 5.3.2.1.5 for further details). By default, this approach makes the assumption that the right-censoring mechanism is homogeneous among all participants, which may not be a valid assumption in this study; censoring in the context of this analysis will come from study termination and it is possible that termination may be higher or lower in the UCB0599 arms compared to placebo. To counter this potential bias, we will obtain treatment group-specific weights by applying the KM estimation method separately to each group (ie, using treatment group as a stratification variable in the model). A summary table presenting the adjusted difference in RMET of ST intake between each UCB0599 treatment group and placebo and the corresponding 95% CIs will be produced, this summary will also include the model estimates of adjusted RMET for each treatment group.

Sensitivity analyses

Additionally, the RMET will be estimated for each combination of treatment group and gender as the area under the KM "survival" curve up to the last observed event time (ie, without using a modeling approach). A plot of these KM curves will be produced, with the RMET estimates also included in the figure. See Section 5.3.2.1.5 for further details.

Supplementary analyses (Cox regression)

As a supplementary analysis a Cox regression model will be fitted to all observed data since this is the traditional way of modelling time-to-event data. This approach assumes proportional hazards for all individuals and the validity of this assumption will be assessed as detailed in the sensitivity analyses section above. Gender and age at baseline will be included as covariates in the model and treatment group as the effect of interest. A summary table presenting the hazard ratios of ST initiation (UCB0599 high dose vs placebo and UCB0599 low dose vs placebo) and the corresponding 95% CIs will be produced.

5.3.2.4.2 Intake of ST

Definitions of endpoints

ST intake will be analyzed at both 12 and 18 months. For both estimands, the participant level endpoint is a binary variable indicating whether the participant has started ST or not, at 12 and 18 months respectively.

Main analytical approach

Full details on the ICE handling for these estimands can be found in Section 1.2. The main ICEs for these estimands will be study termination and loss to follow up (see details in Table 1–3).

At both time points, the population-level summary will be the population odds ratio of ST intake between:

- UCB0599 high dose (360mg/day) and placebo
- UCB0599 low dose (180mg/day) and placebo

In other words, population ratio of odds of ST intake in the presence of UCB0599 over odds of ST intake in the absence of UCB0599.

For both analyses a logistic regression will be performed, with gender and age at Baseline included as covariates. Summary tables will be produced displaying the odds ratios, 95% confidence intervals for the odds ratios, and the p-values for each comparison. This summary will also include the number of participants in each treatment group who are/are not on ST at the timepoint being analyzed.

Sensitivity analyses

A relative risk (RR) regression using a log-link binomial generalized linear model (GLM) will also be fitted to this binary outcome to obtain the RR point estimates and corresponding 95% CIs. This regression will use the same covariates and fixed effects as the logistic regression model. This analysis will be conducted for 12 and 18 months.

5.3.3 Exploratory endpoints analysis

5.3.3.1 Exploratory endpoints: clinical symptoms

All exploratory clinical symptoms endpoints will be summarized descriptively, no statistical analyses are planned for these endpoints within the CSR. All summaries and listings will be presented by visit and treatment group where applicable, listings and summary tables will be produced based on the FAS. All data will be listed by study participant and individual item (where applicable), with a flag for ST initiation at each visit (Yes/No flag for whether they have started taking ST by that visit) included in the listing.

Where age categories are being used in plots, the following categories will be defined (in years):

- 40 to less than 50;
- 50 to less than 60;
- 60 to less than 70;
- 70+ years.

For all continuous endpoints, summary statistics will be produced for both observed results and changes from Baseline.

Patient reported outcomes (PROs)

Fatigue PRO

The Fatigue PRO is composed of 3 scales containing 31 items each with 5 levels: Physical fatigue (items 1-9, raw score range 0-36), Mental fatigue (items 10-20, raw score range 0-44) and Fatigability (items 21-31, raw score range 0-44). Before calculating the sum scores for each fatigue scale, the items for all these scales need to be rescaled from 1-5 to 0-4 (as for a Likert scale). This PRO will be summarized using continuous summary statistics by visit and treatment group. The summary will be presented overall and for participants who are not yet on ST. This summary will be repeated for the transformed score, separately for each scale, where transformed score is calculated as:

$$\frac{raw\ score}{raw\ score\ range} \times \frac{total\ number\ of\ items\ in\ the\ scale}{number\ of\ non-missing\ items\ in\ the\ scale} \times 100$$

Plots of individual trajectories over time will also be produced based on the raw score of each of the 3 fatigue scales. To make these plots interpretable with our large sample size, plots will be produced by treatment group, gender and age category (using the categories defined at the beginning of this section). Trajectories will be color coded to clearly indicate when a participant is and is not on ST. In addition to these individual plots, plots of observed means and mean change from Baseline will be produced for each fatigue scale, on both the raw and transformed scale, by treatment group and ST initiation timing category, overall and by gender. These plots will include error bars (+/- SD) and all treatment groups will be overlaid on the same plot.

Early PD Function Slowness and Early PD Mobility PROs

The early PD function slowness PRO has 28 items each with 5 levels (0 to 4, with 4 representing "slower function") and the early PD mobility PRO consists of 23 items each with 5 levels (0 to 4, with 4 representing "worse mobility"). Data for these PROs will be listed only. The analyses of these PROs will be described in the PD0053 Exploratory Analysis Plan; the results will be part of a separate exploratory report and will not be part of the CSR.

PGIS and PGIC

Patient Global Impression of Severity (PGIS, with 2 items: PD symptoms and fatigue) and Patient Global Impression of Change (PGIC, with 2 items: PD symptoms and fatigue) will be summarized descriptively through frequency counts and percentages of individual response categories by visit and treatment group. The summaries will be presented overall and for participants who are not yet on ST.

A shift table for PGIS will be produced by treatment group and visit, overall and for participants who are not yet on ST at a particular visit. This summary will present a cross-tabulation of Baseline values against post-Baseline values. Each cell will include the number and percentage of participants who have "shifted" between the two categories.

HADS

For the Hospital Anxiety and Depression Scale (HADS), two sub-scores, Depression (7-items, raw score range 0-21) and Anxiety (7 items, raw score range 0-21) will be calculated by summing the respective item scores. Each sub-score will be summarized using continuous summary statistics presented by visit and treatment group. The summaries will be presented overall and for participants who are not yet on ST.

Plots of individual trajectories over time will also be produced. To make these plots interpretable with our large sample size, plots will be produced by treatment group, gender and age category (using the categories defined at the beginning of this section). Trajectories will be color coded to clearly indicate when a participant initiates ST.

In addition, each sub-score which will be categorized as follows:

- 0-7 = Normal;
- 8-10 = Borderline abnormal;
- 11-21 = Abnormal.

The two categorized sub-scores will be summarized descriptively through frequency counts and percentages for individual items presented by visit and treatment group. The results will be presented overall and for participants who are not yet on ST. Shift tables will be produced by treatment group and visit, overall and for participants who are not yet on ST for this categorized version of the HADS.

EQ-5D-5L

The EQ-5D-5L consists of the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS) (EuroQol Group, 2017). The EQ-5D-5L descriptive system comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression scored according to 5 levels: no problems, slight problems, moderate problems, severe problems and extreme problems. These levels are expressed as a 1-digit number, and the digits for the five dimensions are combined into a 5-digit combined score that describes the participants health state (eg, 13414). The 1-digit numbers for each of the 5 dimensions will be summarized categorically by visit and treatment group. Shift tables will also be produced by dimension. The 5-digit health state combined score will be used in the listings.

The EQ VAS records a patient's self-rated health on a vertical visual analogue scale; it ranges from 0 to 100 (with 100 representing the "best health you can imagine" and 0 representing "The worst health you can imagine"). EQ VAS scores will be summarized using continuous descriptive statistics presented by visit and treatment group. For both sets of EQ scores/ratings, data will be summarized overall and for participants who are not yet on ST.

In addition to the summary table, plots of individual trajectories for the EQ VAS score will be produced by treatment group, gender and age category. As done for the plot of HADS, trajectories will be color coded to clearly indicate when a participant initiates ST.

Other exploratory clinical outcomes

Time to first increase in MDS-UPDRS Part I specific items

Any analysis regarding time to first increase in MDS-UPDRS Part I will be done as part of exploratory analyses, outside of the CSR.

Single-item scales

Modified Hoehn and Yahr staging results will be summarized descriptively using frequency counts and percentages. Scores/ratings will be presented overall and for participants who are not yet on ST at a particular visit, by treatment group and visit.

Clinical Global Impressions of Improvement (CGII), Clinical Global Impressions of Severity (CGIS) and Schwab and England Activities of Daily Living (SE-ADL) will all be summarized in the same way as the modified Hoehn and Yahr scores/ratings. Since SE-ADL categorical responses are percentages ranging from 0-100% (11 categories going up in 10% intervals), this will also be summarized as a continuous variable.

In addition to the above summaries, for all of these single-item scales (with the exception of CGII) shift tables will be produced by treatment group and visit, overall and for participants who are not yet on ST at a particular visit. These summaries will present a cross-tabulation of Baseline values against post-Baseline values. Each cell will include the number and percentage of participants who have "shifted" between the two categories. For SE-ADL a 70% threshold will be used, summarizing shifts from 70% or below to above 70%.

Starkstein Apathy Scale

The 14-item Starkstein Apathy Scale (SAS) (using a 4-point Likert scale scoring 0-1-2-3, with the raw score ranging from 0 to 42) will be summarized using continuous summary statistics by visit and treatment group. The summary will be presented overall and for participants who are not yet on ST. A plot of individual trajectories over time will be produced, using the same by variables and color-coding described for the PRO trajectory plots.

Movement Disorder Society -Non-Motor Scale

The MDS-NMS scale is composed of 63 items grouped in 15 domains. Continuous summary statistics will be presented for the MDS-Non-Motor Scale (MDS-NMS) total scores (total frequency and total severity scores) by treatment group and visit. The data listing for MDS-NMS will be grouped by domain, with frequency and severity results appearing side by side. A plot of individual trajectories over time will be produced for the total score, using the same by variables and color-coding described for the PRO trajectory plots.

In addition, the scale includes a new 8-domain non-motor fluctuations subscale, results for this section will not be listed or summarized.

Composite endpoint (based on MDS-UPDRS and/or Early PD PROs)

Another aim of this study is to investigate the use of a new composite endpoint based on MDS-UPDRS and/or Early PD PROs, which may prove a more sensitive endpoint to detect a treatment effect. These analyses are described in a separate analysis plan and will not be part of the CSR.

Wearable sensor

Wearable sensor data is also collected in this study, However, the analyses will not be part of the TFLs and will not be reported in the CSR. A separate report will be produced for these analyses (and potentially added on to the CSR as an appendix).

The wearable sensor also measures MDS-UPDRS Part III items. This data will be analyzed outside of the CSR using the same analysis methods planned in this SAP to allow for comparison of results.

5.3.3.2 Exploratory neurodegeneration endpoints

DaT-SPECT

DaT-SPECT regional SBR continuous summary statistics will be presented by visit and treatment group for the following striatal sub-regions: ipsilateral caudate small, ipsilateral putamen small, contralateral caudate small and contralateral putamen small. Ipsilateral and contralateral side will be derived based on the laterality of motor symptoms dominance at Baseline. If dominant laterality is right then right ROI is defined as ipsilateral and left ROI as contralateral and vice versa. These are the sub-regions used to calculate the whole striatum SBR (listed, summarized and analyzed as a secondary estimand analysis).The SBR for each region of interest will be calculated with the occipital cortex as a reference region using the formula from Section 5.3.2.3. The summaries will be presented overall and for participants who are not yet on ST.

Plots of individual trajectories over time for each striatal sub-region will also be produced, by treatment group, gender and age category. All plots will be color-coded based on ST initiation (as described for plots in Section 5.3.3.1).

In addition to these individual plots, longitudinal plots of mean observed data will be produced by treatment group, overall and by gender. These plots will include error bars (+/- SD) and all treatment groups will be overlaid on the same plot.

CSF total alpha-syn and CSF ASYN oligomers/seeding capacity

CSF total alpha-syn continuous summary statistics will be presented by visit and treatment group, the summary will be presented overall and for participants who are not yet on ST. A trajectory plot, using the same by-variables and color-coding as described for DaT-SPECT mean SBR striatal sub-regions will also be produced. Plots of observed means will be produced by treatment group, overall and by gender. These plots will include error bars (+/- SD) and all treatment groups will be overlaid on the same plot.

Analyses of CSF ASYN oligomers/seeding capacity will be covered in the Exploratory Analysis Plan.

5.3.3.3 Exploratory ST intake endpoints

Cumulative Levodopa equivalent daily dose (LEDD)

The cumulative LEDD will be calculated for each participant at each visit. This is the sum of all the levodopa equivalent daily doses taken up to that visit (ie, if a participant is taking 300mg of levodopa a day, their cumulative dose over 10 days would be 3000mg). A listing of cumulative LEDD will be produced by visit and treatment group based on the FAS, only participants with non-zero data will be presented in this listing. Continuous summary statistics will be presented by visit, treatment group, overall and by gender. Summary statistics will only be calculated at visits 10, 11, 12 and 15 (month 8, 10, 12 and 18). For participants who have not started ST by a particular visit, their cumulative LEDD will be 0mg and this will be included in the calculation of summary statistics. As part of the summary statistics, the number of participants who have initiated ST by each visit will be included. In addition to the summary table, plots of individual trajectories of cumulative LEDD will be produced by treatment group, gender and age category.

ST options other than levodopa are only permitted in exceptional cases. For participants taking levodopa, the LEDD calculation is simple and the cumulative dose is calculated just based on the dose level, dosing regimen and duration of ST. For participants not taking levodopa, the levodopa equivalent dose will be calculated using conversion factors as detailed in the study ADRG (ADaM data reviewers guide).

When calculating the LEDD, the following points should be taken into consideration:

- All participants taking levodopa will take carbidopa (or equivalent) in combination with levodopa, carbidopa intake is ignored when calculating levodopa equivalent dosing. A table of conversion factors for other medications is shown below:
- MAO-B inhibitors are prohibited medications in this study, taking any of these medications would lead to a protocol deviation.
- Despite not being the recommended approach per protocol, participants may be prescribed levodopa in combination with a COMT inhibitor. The COMT inhibitor medication may be prescribed as a separate pill or part of a combination drug with levodopa. COMT inhibitor conversion factors will depend on the dosing level and regimen, the ADRG will give further details on this.

Any changes in medication (type, dose or dosing regimen) should be accounted for when calculating cumulative doses. Participants who take a ST other than levodopa/carbidopa at any point during the study will be flagged in the listing of cumulative LEDD.

5.4 Safety analyses

Safety data will be summarized and listed by treatment group for the SS, with the exception of AE listings which will be presented for the ASPS.

5.4.1 Extent of exposure

Exposure data will be listed for each study participant in the SS by treatment group. This listing will include date of first dose, date of last dose, total number of capsules taken, duration of exposure and overall IMP compliance (see Section 6.1.8 for how this is calculated).

The duration of exposure to IMP (days) will be calculated as follows:

$$Duration = (Date\ of\ Last\ Dose - Date\ of\ First\ Dose) + 1\ days$$

Participants who have any dosing interruptions during the study (temporarily stop taking medication) will have multiple rows in this table. Therefore, a column for "cumulative duration" will also be presented. Participants who have a dosing interruption will be flagged.

Duration of exposure and total number of capsules taken will be summarized using descriptive statistics by treatment group based on the SS and will be presented in the same summary table as study medication compliance (see Section 6.1.8).

5.4.2 Adverse events

AEs will be collected from the time of signing the informed consent form (ICF) until the final SFU visit, and will be characterized as pre-treatment and treatment emergent according to intake of study medication. Adverse events with a start date prior to the first dose of treatment will be defined as pre-treatment AEs. A treatment-emergent AE (TEAE) is defined as any AE with a start date on or after the first dose of treatment or any unresolved event already present before administration of treatment that worsens in intensity following exposure to the treatment. Any AE with onset date later than the SFU visit will not be considered as treatment-emergent and therefore will not be included in the tabulations of TEAEs. Where dates are missing or partially missing, AEs will be assumed to be treatment-emergent, unless there is clear evidence to suggest otherwise. Missing or partially missing dates for AEs will be handled as described in Section 5.1.1.

An overview of the occurrence and incidence of TEAEs will be provided by treatment group. The overview will present individual occurrences as well as number and percentage of participants experiencing any of the following:

- Any TEAE
- Any Serious TEAE
- Participant discontinuation due to TEAEs
- Any drug related TEAEs
- Any severe TEAEs
- Permanent withdrawal of study medication due to TEAEs
- All deaths (AEs leading to death)
- Deaths (TEAEs leading to death)

Summaries of the occurrence and incidence of SAEs and TEAEs will be provided by MedDRA® SOC, HLT, PT and treatment group (including a "UCB0599 Total" column which pools the high and low dose groups). These summaries will be provided for the following:

- Incidence of TEAEs
- Incidence of SAEs
- Incidence of serious TEAEs and serious TEAEs by relationship
- Incidence of fatal TEAEs and fatal TEAEs by relationship
- Incidence of TEAEs leading to discontinuation
- Incidence of non-serious TEAEs
- Incidence of TEAEs by maximum intensity
- Incidence of TEAEs by relationship and maximum relationship
- Incidence of non-serious TEAEs above reporting threshold of 5% of study participants

- Incidence of non-serious TEAEs above reporting threshold of 5% of study participants by relationship

Summary tables will contain counts of study participant, percentage of study participants in parentheses and the number of events where applicable. A participant who has multiple events in the same SOC, HLT and PT during a given treatment will be counted only once in the participant counts for that treatment, but all events will be included.

In summaries including relationship, the following relationships will be summarized: 'Not related' and 'Related'. Participants who experience the same event multiple times will be included in the most related category for tabulations by maximum relationship. Events with missing relationship will be considered as 'Related' for summary purposes but recorded as missing in the listings.

In summaries including intensity, the following intensity categories will be summarized: 'Mild', 'Moderate', 'Severe'. Participants who experience the same event multiple times will be included in the most severe category for tabulations by maximum intensity. Events with missing intensity will be considered as 'Severe' events for summary purposes but recorded as missing in the listings.

Adverse event summaries will be ordered by alphabetical SOC, alphabetical HLT and decreasing frequency of PT within SOC in the "UCB0599 Total" group.

The following listings of AEs will be presented based on the ASPS:

- All AEs
- All Serious AEs
- Discontinuation due to AEs
- AEs leading to death
- AEs related to the wearable sensors
- COVID-19 Infections

All listings will be presented by treatment group and study participant and will include the onset date and outcome date of the event, the event duration (derived), time to onset (derived), pattern of event, intensity, relationship to study medication, action taken and outcome. AEs that led to discontinuation, TEAEs, AESIs and SAEs will be flagged where applicable.

In addition to the above, the following summary tables will be produced for use in the study Plain Language Summary (and are not required for the CSR):

- Incidence of Drug-related TEAEs (Overview)
- Incidence of Drug-related TEAEs by Preferred Term
- Incidence of Serious Drug-related TEAEs by Preferred Term

5.4.2.1 Adverse events of special interest (AESI)

An AESI is any AE that a regulatory authority has mandated be reported on an expedited basis, regardless of the seriousness, expectedness, or relatedness of the AE to the administration of a UCB product/compound.

For UCB0599, the following events require immediate reporting (within 24 hours regardless of seriousness) to UCB:

- Potential Hy's Law, defined as ≥3x upper limit of normal (ULN) alanine aminotransferase (ALT) or aspartate aminotransferase (AST) with coexisting ≥2xULN total bilirubin in the absence of ≥2xULN alkaline phosphatase (ALP), with no alternative explanation for the biochemical abnormality (ie, without waiting for any additional etiologic investigations to have been concluded). Follow up information should then be reported if an alternative etiology is identified during investigation and monitoring of the participant.
- Hypersensitivity reactions (such as rash, angioedema, or anaphylaxis).

These AESIs will be identified based on MedDRA algorithm, standardized MedDRA queries (SMQs), HLTs and PTs as provided in Section 6.1.6 of this SAP. These events will also be identified by the investigator, the investigator flag will be used for flagging events in our listings and for presenting summary tables. Any discrepancies between these two sources will be investigated.

AESIs will be flagged in the AE listings, they will also be presented in a separate listing. A summary of the occurrence and incidence of treatment-emergent AESIs will be provided by MedDRA® SOC, HLT, PT, treatment group and country. This summary will also be repeated by relationship to study medication.

5.4.3 Additional safety assessments

5.4.3.1 Clinical safety laboratory assessments

Hematology, clinical chemistry, coagulation and urinalysis parameters will be assessed, see table 5.3 below for a full list of which parameters will be analyzed. Other screening and laboratory tests to be carried out are also covered in this table. The schedule of activities (protocol section 1.3) gives details on when these assessments are performed.

Hematology, clinical chemistry and coagulation laboratory results that were performed locally should be included in laboratory listings but not in summary tables and figures. Laboratory re-test results will be used in summary tables only if the original scheduled test result is not available, if the original result is available re-tests will be treated as unscheduled.

Table 5–3: Clinical laboratory assessments

Laboratory Assessments	Parameters			
Hematology	Platelet Count RBC Count[b] Hemoglobin[b] Hematocrit	RBC Indices: MCV MCH %Reticulocytes	WBC Count with Differential: Neutrophils[b] Lymphocytes[b] Monocytes Eosinophils[b] Basophils[b]	
Clinical Chemistry	Blood Urea Nitrogen (BUN)	Potassium	Aspartate Aminotransferase (AST)/ Serum Glutamic-Oxaloacetic Transaminase (SGOT) [b]	Total and direct bilirubin
	Creatinine	Sodium	Alanine Aminotransferase (ALT)/ Serum Glutamic-Pyruvic Transaminase (SGPT) [b]	Total Protein
	Glucose (fasting at Screening and nonfasting at any visit thereafter)	Calcium	Alkaline phosphatase[b]	Creatinine phosphokinase, Lactate dehydrogenase, serum aldolase, Cystatin C
Coagulation	International normalized ratio	Prothrombin time[b]	aPTT	Fibrinogen
Routine Urinalysis	• Specific gravity • pH, glucose, protein, blood, ketones, bilirubin, urobilinogen, nitrite, leukocytes by dipstick • Microscopic examination (if blood or protein is abnormal)			
Other Screening Tests	• Follicle-stimulating hormone and estradiol (as needed in women of non-childbearing potential only) • Immunoglobulin E • HbA1c (as needed in study participants with type 2 diabetes mellitus) • Urine drug screen (to include at minimum: amphetamines, barbiturates, cocaine, opiates, cannabinoids, methadone, and benzodiazepines) • Urine alcohol test • Serum and urine human chorionic gonadotropin (hCG) pregnancy test (as needed for women of childbearing potential) [a]			

Table 5–3: Clinical laboratory assessments

Laboratory Assessments	Parameters
	• Serology (HIV antibody, hepatitis B surface antigen, and hepatitis C virus antibody All study-required laboratory assessments will be performed by a central laboratory. The results of each test must be entered into the eCRF.
Other Laboratory Tests	• RBC count and/or hemoglobin in CSF for quality control purposes

Note: Details of liver chemistry stopping criteria and required actions and follow-up assessments after liver stopping or monitoring event are given in Section 7.1.1 of the protocol and Appendix 6 of the protocol (Section 10.6). All events of ALT ≥3×upper limit of normal (ULN) and bilirubin ≥2 ULN (>35% direct bilirubin) or ALT ≥3×ULN and international normalized ratio (INR) >1.5, if INR measured, may indicate severe liver injury (possible Hy's Law) and must be reported as an SAE.

[a] Local urine testing will be standard for the protocol unless serum testing is required by local regulation or IRB.

[b] Shift tables will be presented for these variables.

5.4.3.1.1 Laboratory values over time

Laboratory variables will be grouped according to the laboratory function panel and categorized as normal, high or low, if applicable, based on the reference range supplied by the analytical laboratory.

For urinalysis, hematology, coagulation and clinical chemistry observed results and changes from Baseline for numeric variables at each post-Baseline visit will be listed by treatment group, study participant and visit. Listings will flag values that fall outside of the normal range, and will include the reference ranges. Additionally, values that fall outside the reference ranges will be listed separately.

Clinical chemistry and hematology variables will be summarized descriptively by treatment group at each visit, for both observed values and changes from Baseline. These descriptive statistics will only be calculated if at most one third of the individual data points at any visit are missing or not quantifiable. If 3 or less participants have data available at any visit, only n, minimum and maximum will be presented. Plots of mean and mean change from Baseline will be presented by treatment group for all hematology and clinical chemistry laboratory variables. These plots will include all treatment groups overlaid on the same plot and will include error bars based on the SD (ie, mean +/- SD).

Measurements that are below the limit of quantification (BLQ) or above the limit of quantification (ALQ) will be presented as BLQ and ALQ in the listings. For the purpose of calculating change from Baseline or for descriptive statistics, BLQ values will be imputed with half of the lower limit of quantification (LLOQ) and ALQ values will be imputed to the upper quantification limit (if applicable).

A separate listing will be produced for study participants with Elevated Liver Function Results. Elevated results are defined as results meeting the following criteria:

- ALT or AST ≥3x (ULN)
- ALP ≥1.5xULN
- Total bilirubin ≥1.5xULN

The listing will display all scheduled and unscheduled visit data for participants who meet at least one of the above criteria was fulfilled. A summary of study participants who meet the criteria for potential drug-induced liver injury (PDILI) will be presented and a figure presenting liver function results over time for participants who meet the potential Hy's Law Criteria will be produced (this figure will only be produced if there are participants who meet the criteria).

Four additional outputs for liver function tests will also be produced:

- A summary table for elevated liver function tests
- A figure showing shifts in liver functions tests from baseline to maximum post-baseline results
- A figure showing Maximum post-baseline total bilirubin versus maximum post-baseline ALT
- A figure presenting the liver function results over time for participants who meet the PDILI criteria (participants with ALT or AST>=3xULN).

Any additional laboratory variables not included in the outputs described previously will be listed separately.

5.4.3.1.2 Individual participant changes of laboratory values

For selected variables that are identified in Table 5–3, the change in category from Baseline will be presented in shift tables for all post-Baseline visits. These summaries will present a cross-tabulation of Baseline values against post-Baseline values categorized as below normal range, within normal range or above normal range. Each cell will include the number and percentage of participants who have "shifted" between the two categories.

5.4.3.2 Vital signs

The following vital signs measurements will be assessed at every visit throughout the study, after 5 minutes of rest in the supine position and erect (to assess autonomous dysregulation):

- Systolic and diastolic blood pressure (3 readings, all readings will be recorded in the CRF and the average will be derived for analyses)
- Pulse rate (3 readings, all readings will be recorded in the CRF and the average will be derived for analyses)
- Tympanic body temperature (1 reading)
- Respiratory rate (1 reading)

All vital signs results will be listed by treatment group, study participant and visit. For systolic blood pressure, diastolic blood pressure and pulse rate all 3 readings as well as the average will be including in the listing. The listing will include observed results, change from Baseline and a flag for abnormal values. For assessments taken 3 times, change from baseline will only be calculated and listed for the average of the 3 readings.

5.4.3.2.1 Vital sign values over time

Vital signs measurements (observed values and changes from Baseline) will be summarized by treatment group, measurement, position and visit. For assessments taken 3 times, only the average of the 3 readings will be used in summary tables.

The number and percentage of study participants with treatment-emergent markedly abnormal (TEMA)/potentially clinically significant (PCS) vital sign values as calculated by the criteria outlined in the table below will be summarized by treatment group and visit.

Table 5–4: TEMA/PCS criteria for vital signs

Variable	Unit	Low[a]	High[a]
Systolic blood pressure	mmHg	Value ≤90 and ≥20 decrease from Baseline	Value ≥180 and ≥20 increase from Baseline
Diastolic blood pressure	mmHg	Value ≤50 and ≥15 decrease from Baseline	Value ≥105 and ≥15 increase from Baseline
Pulse rate	bpm	Value ≤50 and ≥15 decrease from Baseline	Value ≥120 and ≥15 increase from Baseline
Respiratory rate	Breaths per minute	<12 and decrease of ≥5	>20 and increase of ≥5

bpm=beats per minute; PCS=potentially clinically significant; TEMA=treatment-emergent markedly abnormal.

[a] Both conditions must be satisfied for a measurement to be considered potentially clinically significant.

5.4.3.3 Electrocardiograms

All standard 12-lead ECG recordings will be taken in triplicate (as closely as possible in succession but no more than 2 minutes apart) with the study participant resting in the supine position for at least 10 minutes before recording. The individual mean at each time point will be calculated, all summary tables will be based on the mean. In the event that there are not 3 available measurements at any given time, the mean will be calculated based on the number of measurements for which data are provided.

The following ECG parameters will be reported:

- Heart rate (bpm)
- PR interval (ms)
- QRS duration (ms)
- QT interval (ms)
- QTcF interval (QT corrected for heart rate using Fridericia's formula) (ms)
- Investigator's conclusion on ECG profile

5.4.3.3.1 Electrocardiogram values over time

Individual measurements and the mean of the triplicate measurements will be listed. This listing will include change from Baseline (based on the mean of the triplicate measurements) and will be presented by treatment group and visit.

Observed values and change from Baseline will be summarized by treatment group, ECG variable and visit (based on the mean of the triplicate values at each visit). The mean change from Baseline and its 95% CI for each ECG parameter will also be summarized graphically over scheduled time points with all treatment groups overlaid on the same plot.

The following cut-points in QTcF based on the mean of the triplicate data will be summarized categorically (number and percentage of participants) by treatment group and visit.

For observed data:

- <450 msec (milliseconds)
- ≥450 to <480 msec
- ≥480 to <500 msec
- ≥500 msec

For change from Baseline in QTcF:

- <30 msec
- ≥30 to <60 msec
- ≥60 msec

All ECG findings for the individual triplicate measurements will be listed separately.

5.4.3.4 Physical and neurological examination

Study participants with abnormalities in the physical and neurological examination will have this information listed including details of the abnormality.

5.4.3.5 Suicidal risk monitoring

The Columbia-Suicide Severity Rating Scale (C-SSRS; Posner et al, 2011) evaluations include suicidal ideation, intensity of ideation, suicidal behavior, and actual attempts and will be performed at all study visits (except for screening visit 2). Module of the questionnaire, time point, question and the associated response will be listed for all visits where this questionnaire is collected by treatment group. Only data where suicidal ideation/behaviour has been reported needs to be included in the listing.

5.4.3.6 MRI assessments

Structural brain MRIs will be performed at the Screening Visit (Visit 2) for study participants without an acceptable historical scan within 6 months before this visit. The date of these scans will be available in the clinical database, but no information from these MRIs will be presented in the TFLs.

5.4.3.7 Wearable sensor data

Physiological and environmental data obtained from the wearable sensor will not be used for safety monitoring and reporting of UCB0599 because the sensor is unvalidated for this purpose and will be only used for exploratory endpoints in this study.

The Investigator is expected to record any AEs in the eCRF and assess causality as it might relate to the wearable sensor. A listing of these AEs will be produced for the ASPS.

All other analyses will be done outside of the CSR and a separate report will be produced.

5.5 Other analyses

5.5.1 Pharmacokinetics

Individual plasma and CSF concentrations of UCB0599 and its N-oxide metabolites will be listed separately for the SS and will include the actual sampling times. Any samples that are obtained outside the tolerance window permitted at the specified time point will be discussed at a DEM.

Summary analyses will be performed on the PKS by treatment group. Concentration data (plasma and CSF) for both UCB0599 and its N-oxide metabolites will be summarized separately by nominal sampling times using descriptive statistics. Descriptive statistics to be presented are: n, arithmetic mean, median, SD, minimum, maximum, geometric mean, geometric CV and 95% CI for the geometric mean (assuming log-normally distributed data).

Combined individual concentration time-profiles (spaghetti plots) for the PKS will be produced, with all participants overlaid on the same plot. Geometric mean profiles over time will be presented, on both linear and semi-logarithmic scales (geometric means will be calculated by treatment group and overlaid on the same plot). The 95% CI for the geometric mean will be displayed for the linear scale plot only. All figures will be produced for both plasma concentration and CSF concentration data, for both UCB0599 and its N-oxide metabolite by treatment group.

CSF:plasma concentration ratios will be calculated for each time point where both plasma and CSF samples are collected. This data will be listed by treatment group and participant based on the SS, summary tables of descriptive statistics for the PKS will also be produced. Spaghetti plots of these ratios over time will be presented by treatment group, with the mean value at each visit overlaid.

The following rules apply for concentration data listings and summaries:

- Individual concentration data will be reported to the same level of precision as received from the bioanalytical lab
- Values below the LLOQ will be reported as BLQ in the listings
- Descriptive statistics for concentration data will be reported to the same level of precision as the individual data for the minimum and maximum, and to 1 additional decimal place (or 1 additional significant figure depending on the format of the lab data) for the arithmetic mean, geometric mean, median and SD. The 95% CI for the geometric mean will use 1 additional decimal place (or significant figure) compared to the geometric mean.
- Descriptive statistics of concentration data will be calculated if at most 1/3rd of data points at a given time point are missing or not quantifiable (<LLOQ). If there is enough quantifiable data, BLQ values will be replaced by LLOQ/2 for the purposes of calculating summary statistics. If n is less than 3 at any visit, only n, minimum and maximum will be presented, if n=3 only n, minimum, maximum and median will be presented. If no participants have data, only n=0 will be presented.
- Predose concentration data that is confirmed to not have been collected prior to dosing (the date of sample collection and medication administration is the same, but the sample time is after medication time) will be included in the listing but not in the summary tables or figures.
- The 95% CI limits should be left blank if the SD (or equivalently, the geometric CV) is 0.
- The geometric CV will be calculated using the following formula where SD is the standard deviation from the log-transformed data, and will be reported as a percentage to 1 decimal place:

$$Geometric\ CV\ (\%) = \sqrt{e^{SD^2} - 1} \times 100$$

5.5.2 Pharmacodynamics

Not applicable.

5.5.3 Biomarkers

Genetic blood sampling will be mandatory (where allowed) for study participation, however only approximately 50 participants per treatment group will undergo CSF sampling. The following samples will be collected at the time points specified in the schedule of activities (protocol section 1.3):

- Blood samples for DNA (genetic biomarkers)
- Blood samples for RNA (genomic biomarkers)
- CSF and blood samples for other biomarkers

A listing will be produced of the blood sample collection times for the RS. Analyses of this biomarker data are exploratory and will not be reported in the CSR, the only biomarker data that will be summarized as part of the CSR is total CSF ASYN (further details given in Section 5.3.3.2). These samples will only be used to further understanding of PD and/or how biomarkers, including genetic variation, may affect response or be affected by treatment with UCB0599 in the treatment of PD.

5.6 Subgroup analyses

No subgroup analyses will be performed.

Selected analyses may be repeated among participants who did not experience tremor at baseline as part of the exploratory analysis plan.

5.7 Interim Analyses

No interim analysis will be conducted, see Section 5.8 for details of the planned analysis once all participants have completed their 12-month visit.

5.8 Planned analysis at 12 months

The primary analysis for the primary efficacy estimand (MDS-UPDRS Part I-III sum score) uses data from the first 12 months of the study; this analysis as well as the analyses for the secondary estimands at 12 months (where applicable) will be carried out at study end (see details in Protocol Section 9.3).

A limited team of independent pharmacometricians and statistical programmers with access to the 12-month data will be unblinded to prepare for the exposure-response analysis to be finalized following completion of the study. Details of unblinding will be described in the unblinding charter, and care will be taken to ensure that the blinded study team remains blinded until database lock. Details of the exposure-response analysis will be provided in a separate analysis plan. The results of the final exposure-response evaluation will be provided outside of the clinical study report at study finalization.

5.9 Data Monitoring Committee (DMC) or other review board

An independent DMC will conduct safety interim reviews of all available unblinded safety data at the following time points:

- When safety data for 3 months of treatment (ie, until Visit 7) are available for 100 participants
- When safety data for 3 months of treatment (ie, until Visit 7) are available for at least 50 participants per arm
- After at least 50% of the participants have been treated for 8 months (ie, approximately 75 participants per arm)
- After at least 50% of the participants have been treated for 14 months (ie, approximately 75 participants per arm)
- Ad hoc as required.

The DMC will provide a recommendation on the continuation of the study. The activities of the DMC will be described in a separate charter.

An SMC will review (approximately every 3 months, with the option to adapt the frequency based on recruitment rates) the available blinded safety data at the following time points:

- After 5 participants/arm have been treated for 2 weeks
- After 25 participants/arm have been treated for 1 month
- After 50 participants/arm have been treated for 1 month
- After 75 participants/arm have been treated for 1 month
- After 100 participants/arm have been treated for 1 month
- Subsequent SMC meetings will occur every 3 months

Details will be described in a separate charter.

6 SUPPORTING DOCUMENTATION

6.1 Appendix 1: Non-key analysis specifications

6.1.1 Baseline characteristics and demographics

6.1.1.1 Demographics

A listing of demographic characteristics will be presented for all study participants by treatment group, based on the ASPS. This will include year of birth, age (in years), sex, country, race, ethnicity, height (in cm), weight (in kg) and body mass index (BMI). The body weight included in this listing (and in all other demographic outputs) will be the value measured at screening visit 1.

Body mass index in kg/m^2 is calculated based on the height (in m) and the weight (in kg) using the following formula (If height is in cm, then height will be converted to meters by dividing by 100):

$$BMI\ (kg/m^2) = weight\ (kg) / [height\ (m)]^2$$

The BMI will be reported to 1 decimal place and should be recalculated even if reported in the CRF.

All demographic characteristics (except for year of birth) will be summarized by treatment group and for all study participants based on the RS and the SS (the summary will only be produced twice if the RS and the SS differ). The summary of age will include descriptive statistics and categorical summaries, the latter based on requirements for European Union Drug Regulating Authorities Clinical Trials (EudraCT) and clinicaltrials.gov reporting.

For the EudraCT reporting, the categories will include:

- 18 to <65 years
- 65 to <85 years
- ≥ 85 years

For the clinicaltrials.gov reporting, the categories will include:

- ≤18 years
- 19 to <65 years
- ≥65 years

6.1.1.2 Baseline disease characteristics

The following Baseline disease characteristics will be summarized by treatment group for the SS:

- MDS-UPDRS Part I-III sum score, Part I, Part II and Part III score
- Modified Hoehn and Yahr Stage
- MoCA

- Duration of disease, will be calculated as follows:

$$Duration\ of\ disease\ (months) = \frac{(Date\ of\ first\ dose\ of\ IMP - Date\ of\ First\ Diagnosis + 1)}{30.3}$$

See Section 5.1.1 for disease duration imputation rules.

For MDS-UPDRS, modified Hoehn and Yahr and MoCA, Day 1 (baseline) data will be used in this summary where available. If Day 1 data is not available, the latest pre-baseline data will be used. This data will also be listed based on the RS, by treatment group and participant.

6.1.1.3 Other baseline characteristics

Not applicable.

6.1.2 Protocol deviations

Important protocol deviations will be identified and classified by the deviation types in the IPD document, further details on how IPDs are defined and identified are given in Section 5.1.2.2.

A listing of all IPDs identified at the DEM will be presented for all participants based on the randomized set and will include the deviation type and description. The number and percentage of participants in the RS with IPDs will be summarized by treatment group and overall, for each deviation type.

6.1.3 Medical history

Medical history and ongoing medical conditions will be listed and summarized (based on the RS and SS respectively) by treatment group, MedDRA® system organ class (SOC) and preferred term (PT). The reported term, start date and stop date will be included in the listing. The summary will include the number and percentage of study participants and will be sorted alphabetically by SOC and by descending incidence of PT within each SOC, based on the 'All Participants' column. History of PD will be considered as a Baseline characteristic and will be summarized separately (Section 6.1.1.2).

Procedure history will be listed by treatment group and study participant for the RS. Family medical history will be collected for any participants with potential drug-induced liver injury, this data will be listed for the SS.

6.1.4 Prior/concomitant medications

Prior medications

Prior medications include any medications that started prior to the date of first dose of study medication. This includes medications that started prior to the first dose and continued after.

Concomitant medications

Concomitant medications included medications with a start date between first (inclusive) dose of study medication and 3 days after the last dose of study medication (inclusive). Medications that started prior to the first dose but stopped after the first dose will be classified as both prior and concomitant.

Prior and concomitant medications will be listed for the RS and summarized (separately) for the SS by treatment group and study participant, and will include WHODD Anatomical Main Group (Level 1), Pharmacological Subgroup (Level 3 term text) and PT. The reported term will be included in the listing. Prior medications which continued into the study period will also be classified as concomitant and will be included in both summaries. And medications with partially missing dates will be handled as described in Section 5.1.1 for prior/concomitant classification.

All tabulations will be sorted alphabetically by Level 1 term, alphabetical Level 3 term within Level 1 and decreasing frequency of PT in the 'All Participants' column.

Since ST intake is an efficacy variable in this study, separate listings and summaries will be made for these medications. These are covered in Section 5.3.2.4.

Concomitant medical procedures will be listed by treatment group and study participant for the RS. Additionally, a separate listing of concomitant COVID-19 vaccinations will be produced so that participants who are vaccinated during the study can be easily identified.

6.1.5 Data derivation rules

Not applicable. Any derived variables are defined as part of the analyses in Section 5.

6.1.6 AEs of special interest

The events defined as AESI for UCB0599 in Section 5.4.2.1 will be summarized as followed:

- Hypersensitivity reactions (such as rash, angioedema, or anaphylaxis)
- Hepatic events and drug induced liver injury.

Hypersensitivity reactions

These will include events based on the following SMQs:

- 'Hypersensitivity' (Narrow)
- 'Severe cutaneous adverse reaction' (Narrow).

Hepatic events and drug induced liver injury

Hepatic events will include:

- Events based on the SMQ = 'Drug related hepatic disorders – comprehensive search' (excluding sub SMQs = 'Liver neoplasms, benign [incl cysts and polyps]' and 'Liver neoplasms, malignant and unspecified'). All AEs should be included in the tabulation (included those considered both related and not related to the IMP) which code to a PT included in the Scope=Narrow group within each SMQ

- Hy's Law cases will also be summarized separately in a table of liver function abnormalities (with adjudication for PDILI cases)

6.1.7 Potentially clinically significant criteria for safety endpoints

Not applicable. All clinically significant criteria for safety endpoints are described in their respective sections.

6.1.8 Compliance

At each in-clinic visit after study medication is dispensed, participants must return all unused study medication and empty study medication containers. Drug accountability must be done in the participant's presence in order to obtain explanations regarding discrepancies in compliance with the dosing regimen. The drug accountability eCRF form collects the amount of study medication dispensed and returned, this will be used to determine whether or not a participant has taken the correct amount. Compliance will be calculated based on the total number of capsules as follows:

$$Compliance\ (\%) = (actual\ number\ of\ doses\ taken\ /\ planned\ number\ of\ doses) \times 100$$

Compliance data will be summarized for the SS, both continuous and categorical summary statistics (by treatment group) will be presented. For categorical summaries, the number and percentage of participants who fall into the following categories will be presented:

- <80% compliant
- 80% to 120% compliant (inclusive)
- >120% compliant

This table will also present summary statistics for duration of exposure (see Section 5.4.1 for details).

In addition to the above calculations for compliance, the listing and summary of compliance data will also present compliance calculated under the assumption that no overdosing has occurred. In this case, when the actual number of doses taken by a participant is greater than the planned number of doses, compliance will be set to 100%. With this calculation we are assuming that the discrepancy between planned and actual dosing is due to drug accountability errors by the participant (for example, lost medication) and not due to the participant taking more medication than needed. All cases where the compliance is set to 100% will be flagged in the data listings.

6.2 Appendix 2: Changes to protocol-planned analyses

The following changes to the protocol-planned analyses have been made in this SAP:

- The definition of the FAS was updated to exclude participants who did not meet key inclusion/exclusion criteria.
- Two new supplementary estimands were added for the primary endpoint, MDS-UPDRS Part I-III sum score. The population-level summaries of interest for these new estimands will be the ST de-mediated difference in target population (1) mean sum score at Month 12 and (2) mean slope over 12 months, as if the participants had not initiated ST.
- Four new supplementary estimands were added for MDS-UPDRS Part III subscale as if the participants had not initiated ST. The population-level summaries of interest for these new estimands will be the ST de-mediated difference in target population (1) mean subscale at Month 12, (2) mean subscale at Month 18, (3) mean slope over 12 months and (4) mean slope over 18 months.
- Four new supplementary estimands were added for MDS-UPDRS Part III ePD subscore as if the participants had not initiated ST. The population-level summaries of interest for these new estimands will be the ST de-mediated difference in target population (1) mean subscale at Month 12, (2) mean subscale at Month 18, (3) mean slope over 12 months and (4) mean slope over 18 months.
- ~~In the protocol, estimands for MDS-UPDRS Part I at Month 12 and Month 18 regardless of ST initiation are considered secondary estimands. In this SAP, they are considered as supplementary estimands and covered in Section 5.3.2.1.8.~~
- For the analysis of MDS-UPDRS Part II emerging symptoms, the ICE handling strategy has been updated from treatment policy to composite for ST for the main analysis, and the treatment policy approach is being proposed as a supplementary analysis.

6.3 Appendix 3: Data loss to ST initiation modelling

Data loss to ST initiation over 12 months as observed for PPMI (Figure 6–1) and assuming exponential loss for our POC study (Figure 6–2).

Figure 6–1: Observed data loss in PPMI dataset

NM

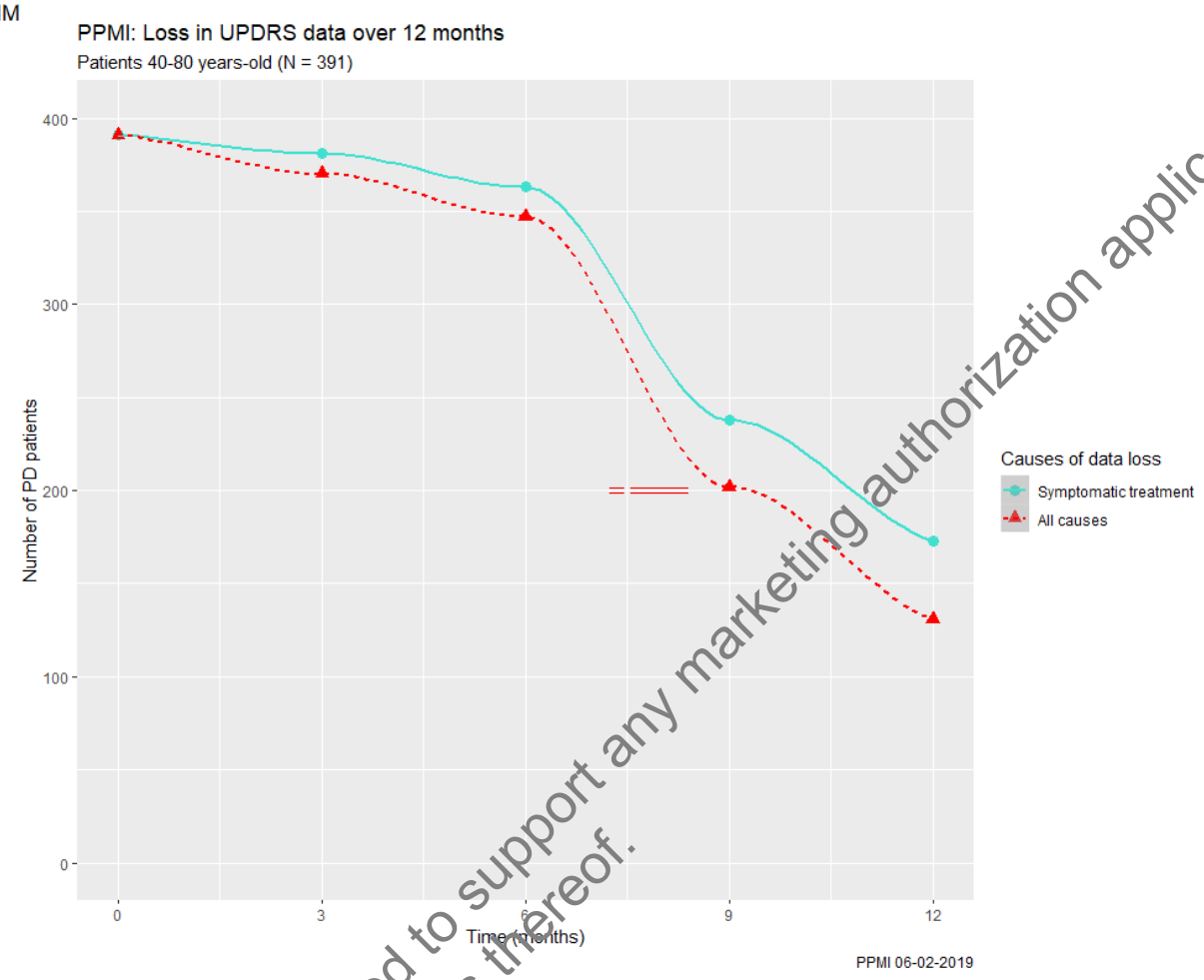
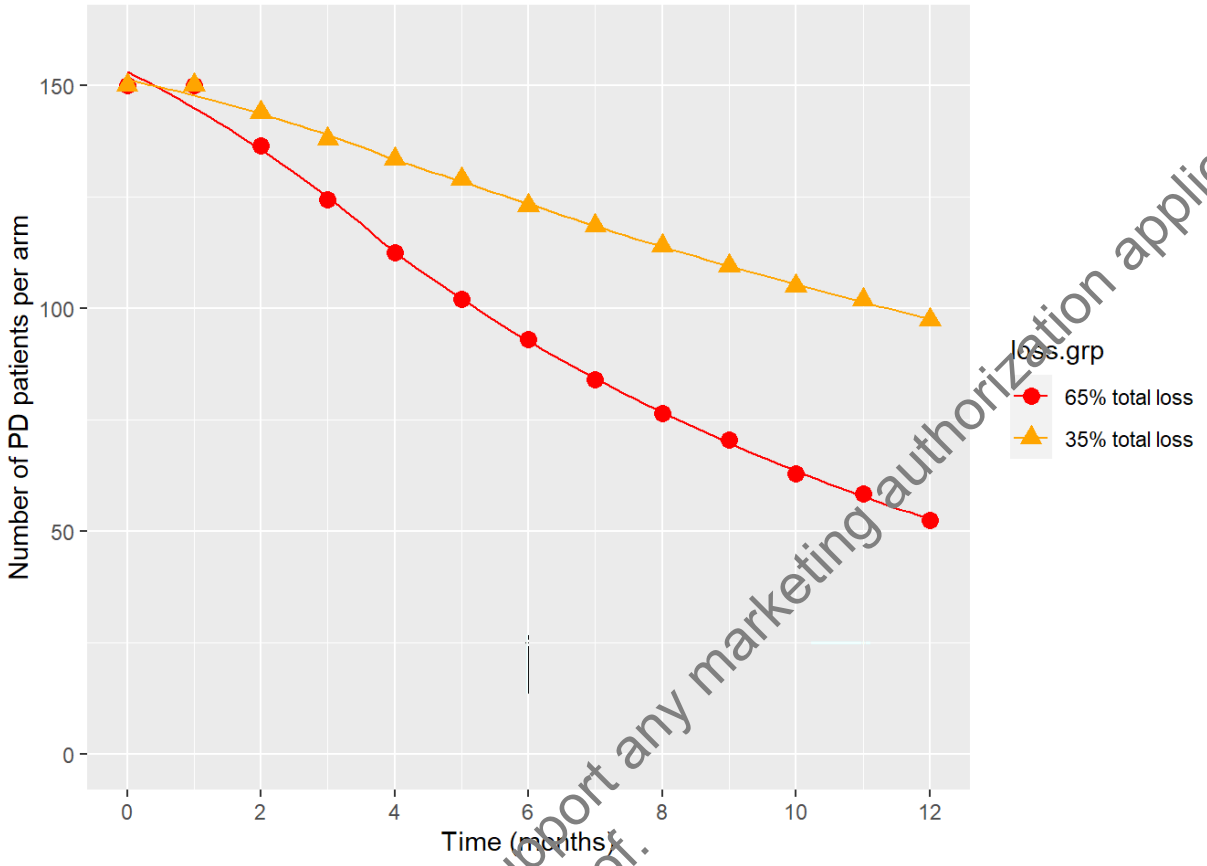


Figure 6–2: Simulated data loss in PD0053



6.4 Appendix 4: Producing summaries by treatment group and overall

Table 6–1: Treatment groups to be presented in summary tables

	Not Randomized	Placebo	UCB0599 Arms	UCB0599 Total	All Participants
Participant disposition	X[a]	X	X		X
Protocol deviations		X	X		X
Demographics and Baseline characteristics		X	X		X
Medical history and medications		X	X		X
Adverse Events		X	X	X	
Other safety analyses		X	X		
IMP Compliance		X	X		
Pharmacokinetics			X		
Efficacy endpoints		X	X		

[a] Only reasons for screen failure and disposition of analysis sets will include not randomized participants.

Note: This table is a guide, the study TFL shells should be followed for which groups to display for each individual output.

6.5 Appendix 5: MDS-UPDRS Part I-III sum score, MDS-UPDRS Part III subscale and ePD subscore: imputation, ST de-mediation and modelling

This Appendix describes in detail the approaches and workflow to be implemented by the UCB statistical/statistical programming team to produce additional (CSR-ready) TFLs for the Month 18 delivery.

6.5.1 Overview of Estimands and General Approach

Estimands

The treatment differences of interest will be:

- Minzasolmin 360mg/day vs Placebo,
- Minzasolmin 180mg/day vs Placebo.

The approaches described below concerns the production of TFLs regarding the following 5 estimands based on 3 MDS-UPDRS based endpoints:

- Symptomatic Treatment (ST) de-mediated difference in MDS-UPDRS Part I-III sum score mean point estimate at Month 12 (section 5.3.1.4.1);
- ST de-mediated difference in MDS-UPDRS Part III subscale mean point estimate at Month 12 (section 5.3.2.1.4);
- ST de-mediated difference in MDS-UPDRS Part III subscale mean point estimate at Month 18 (section 5.3.2.1.4);
- ST de-mediated difference in MDS-UPDRS Part III ePD subscore mean point estimate at Month 12 (section 5.3.2.1.4);
- ST de-mediated difference in MDS-UPDRS Part III ePD subscore mean point estimate at Month 18 (section 5.3.2.1.4).

Five supplementary estimands based on the slope for these 3 endpoints will also be implemented:

- ST de-mediated difference in MDS-UPDRS Part I-III sum score mean slope over 12 months (section 5.3.1.4.3);
- ST de-mediated difference in MDS-UPDRS Part III subscale mean slope over 12 months (section 5.3.2.1.4);
- ST de-mediated difference in MDS-UPDRS Part III subscale mean slope over 18 months (section 5.3.2.1.4);
- ST de-mediated difference in MDS-UPDRS Part III ePD subscore mean slope over 12 months (section 5.3.2.1.4);
- ST de-mediated difference in MDS-UPDRS Part III ePD subscore mean slope over 18 months (section 5.3.2.1.4).

Methodological Approach

ST de-mediation of randomized treatment, to estimate the direct effect of the interventional UCB0599 treatment on the outcome compared to the reference (Placebo) treatment, will be achieved by implementation of a g-estimation approach, which is based on a linear structural nested mean model (SNMM, Vansteelandt & Joffe, 2014).

On implementing the g-estimation, a vector of (MDS-UPDRS) outcomes that would have been seen if no ST had been given is obtained, and the effect of the randomized treatment can be obtained from this vector of ST de-mediated outcomes, by fitting mixed effect models (MEMs) to compare the mean at a given time point (Month 12 or 18), or the slope over a period of time (12 or 18 months) between randomized arms.

Importantly, the standard error for the point estimates must be corrected for the fact that the final analysis model is run on 'predicted' outcomes, i.e. 'predicted' from the SNMMs (case of uncongeniality to the analysis model), which carry uncertainty in the estimated effect of ST. This will be achieved by using a non-parametric bootstrapping approach.

In addition, implementation of the g-estimation requires that the missing data be imputed *a priori.* Intermittent missing data will be imputed assuming missing at random (MAR). Longitudinal/monotone missing data from participants who discontinued the study for reasons unrelated to randomized treatment or who are lost to follow-up (due to unknown reasons by definition) will be imputed assuming MAR, while longitudinal/monotone missing data from participants who discontinued the study for reasons related to randomized treatment (i.e. due to lack of efficacy) will be imputed assuming missing not at random (MNAR), using a relatively simple Reference-Based Imputation approach (RBI), i.e. control-based pattern imputation. As the 3 endpoints of interest are continuous, we will adopt multiple imputation (MI) approaches.

The application of reference-based imputation (RBI) for some of the MNAR data also represents a case of uncongeniality to the analysis model: the imputation model is uncongenial with the analysis model (Bartlett & Hughes, 2020). Therefore, to obtain adequate estimates of standard errors for the point estimates of interest, the recommended approach is to first bootstrap the trial data, then impute via MI. In our case, we will adopt the so-called MI Boot Percentile approach as implemented by (section 3.2.1 in Bartlett & Hughes, 2020), where first B = 200 bootstrap samples of the data are generated, then M = 10 imputations of each B bootstrap sample are generated to obtain B x M datasets for each endpoint. G-estimation models are then applied to each of the B x M dataset, followed by the analysis model(s) and the relevant parameters mean and corresponding (robust) 95% boostrap percentile-CI are then obtained through the approach outlined below. Note that to be conservative, RBI data from participants who discontinued the study for reasons related to randomized treatment will not be ST de-mediated.

Schematic of the Methodological Approach

1. Bootstrap observed data **B** times, by randomized treatment group (described in Section 6.5.3)
2. For each **B** bootstrap sample, use MI to impute missing data **M** times, according to MAR or MNAR paradigms (described in Section 6.5.4)
3. Apply g-estimation once to each **B x M** dataset (described in section 6.5.5), excluding participants who discontinued the study for reasons related to randomized treatment
4. Apply analysis model to each ST de-mediated **B x M** dataset (detailed in section 6.5.6)
5. Obtain dataset-specific mean point estimate and corresponding 95% CI based on bootstrap percentiles over all **B x M** datasets (detailed in section 6.5.7).

6.5.2 Data Pre-processing

Participant selection
See details in section 6.6.

Endpoint datasets

Create 3 separate datasets for the 3 endpoints

- MDS-UPDRS Part I-III sum score from baseline to Month 18;
- MDS-UPDRS Part III subscale from baseline to Month 18;
- MDS-UPDRS Part III ePD subscore from baseline to Month 18.

If a particular subscale (Part I, II or III) is missing at baseline, then the corresponding data from any available visit prior to randomization will be used instead, i.e. most often data collected at the screening visit (section 5.1.2.1.5). Participants with missing baseline data for a particular endpoint will be excluded from the corresponding dataset (section 4).

A description of the items which are part of the MDS-UPDRS Part III subscale are presented in Table 6.2. All items are originally scored on a 5-point Likert scale: '0: No problems', '1: Slight', '2: Mild', '3: Moderate', and '4: Severe'. The ePD subscore selected is the 15-item set, which contains all 5 Rigidity items (3.3), Finger tapping (3.4, Right/Left), Hand movements (3.5, Right/Left), Pronation-supination of the hands (3.6, Right/Left), Toe tapping (3.7, Right/Left), and Leg agility (3.8, Right/Left). The ePD subscore will be constructed by summing the score for the 15 items.

Table 6–2: Items to be included in the 15-item MDS-UPDRS Part III ePD subscore

Item	Description	15-item set
3.1	Speech	
3.2	Facial expression	
3.3	Rigidity in the joints	
	Neck	
	Upper joints/extremities (Right/Left)	
	Lower joints/extremities (Right/Left)	
3.4	Finger tapping (Right/Left)	
3.5	Hand movements (Right/Left)	
3.6	Pronation-supination movements of hands (Right/Left)	
3.7	Toe tapping (Right/Left)	
3.8	Leg agility (Right/Left)	
3.9	Arising from chair	
3.10	Gait	
3.11	Freezing of gait	
3.12	Postural stability	
3.13	Posture	
3.14	Global spontaneity of movement	

Table 6–2: Items to be included in the 15-item MDS-UPDRS Part III ePD subscore

Item	Description	15-item set
3.15	Postural tremor (amplitude) of the hands (Right/Left)	
3.16	Kinetic tremor (amplitude) of the hands (Right/Left)	
3.17	Rest tremor amplitude	
	Lip/Jaw	
	Upper limb/extremity (Right/Left)	
	Lower limb/extremity (Right/Left)	
3.18	Constancy of rest tremor	

Steps described in sections 6.5.3 to 6.5.7 below will be performed for each of the 3 datasets separately.
For all relevant procedures, the seed will be set to **2024**, at the level of the SAS procedure statement.
Validation/QC approaches are detailed in section 6.5.8.

6.5.3 Bootstrapping

- Prior to bootstrapping and in preparation for imputation of longitudinal/monotone missing data (see section 6.5.4), participants will be classified as having monotonic missing data either for reason not related to randomized treatment, or for reasons related to randomized treatment. Participants who discontinued randomized treatment for lack of efficacy prior to study termination or loss-to-follow-up will be included in the latter case.
- Sampling with replacement of each endpoint dataset
- Balanced sampling: sample for each randomized treatment arm separately (STRATA statement)
- Bootstrap **B = 200** times (using **B** as specified in the so-called MI Boot Percentile approach implemented by Bartlett & Hughes, 2020 section 3.2.1)
- Bootstrap using SAS PROC SURVEYSELECT procedure with options METHOD=URS, OUTHITS and SAMPRATE=1.

6.5.4 Handling of Missing Data

Intermittent missing data, i.e., missing values for a given subject that has available data before and after the missing time point, will always be assumed to be MAR.

Longitudinal/monotone missing data, i.e., where all subject data is missing after a given time point, can fall into 2 categories, MAR or MNAR (also refer to section 6.5.3).

Briefly, longitudinal/monotone missing data from participants who discontinued the study for reasons unrelated to randomized treatment or who are lost to follow-up (due to unknown reasons by definition) will be imputed assuming MAR; longitudinal/monotone missing data from participants who discontinued the study for reasons related to randomized treatment (i.e. due to

lack of efficacy) will be imputed assuming missing MNAR, using RBI, here control-based pattern imputation (LingLing, 2019).

Imputation will be applied separately to each **B** bootstrap data samples. A set of **M = 10** imputations per bootstrap data sample will be generated.

For all imputation approaches, time will be set to planned (rather than actual) visit time, i.e. 0, 2, 4, 6, 8, 10, 12, 14, 16 and 18 months, as actual visit time since baseline visit cannot be derived for participants with missing longitudinal/monotone data. Month 18 visit will be set as reference.

Imputation of intermittent MAR data will be implemented using multiple imputation (MI) via Markov Chain Monte Carlo (MCMC) and applying monotone regression (MI-MCMC using the SAS PROC MI procedure with methods/statements MCMC, VAR, with options IMPUTE = monotone and CHAIN = multiple).

Imputation of longitudinal/monotone MAR data will be implemented using MI-MCMC and applying monotone regression (SAS PROC MI procedure with methods/statements MCMC, VAR and MONOTONE REG)

Imputation of longitudinal/monotone MNAR data will be implemented using MI-MCMC, and applying monotone regression but specifying that only observations collected under placebo are to be used to derive the imputation, i.e. RBI (SAS PROC MI procedure with methods/statements MONOTONE REG and MNAR = model).

- For each **B** bootstrap sample, subsets will be created according to each randomized treatment arm, i.e. 3 subsets.
    - Symptomatic treatment (ST) status will be ignored to avoid creating groups with too low sample size for reliable MI.
- MI will be applied to each bootstrap sample subsets separately.
- For each **B x 3** bootstrap sample subset:
    - And for ALL participants, the intermittent missing values will be filled in using the MI-MCMC method, with a total of **M=10** sets of imputations being performed, using data from the same randomized treatment arm subset;
    - For participants randomized to active treatment with monotone MAR data or participants randomized to Placebo, missing data will be imputed by applying monotone regression on these 10 intermittent missing-imputed subsets, using data from the same randomized treatment arm subset;
    - For participants randomized to active treatment with monotone MNAR data, missing data will be imputed by applying monotone regression on these 10 intermittent missing-imputed subsets, using data from the randomized placebo arm subset. If any issues are encountered using the monotone MI assuming MNAR approach, a monotone MI assuming MAR approach will be used instead.

  Imputed dataset from participants with monotone MAR and MNAR data will be stacked back together to reconstitute the full list of trial participants for each imputed set.

- For each **B** bootstrap sample, the **M-**imputed randomized treatment arm subsets will be stacked again into **M** complete imputed datasets corresponding to a single bootstrap sample (full trial).
- If values outside of the pre-defined range of values for MDS-UPDRS endpoints (defined as the minimum and maximum values which could be observed) are imputed, they will be reset to the relevant lower or upper limit of the range after completion of the MI procedure.
    - MDS-UPDRS Part I-III sum score: 0 – 236
    - MDS-UPDRS Part III subscale: 0 – 132
    - MDS-UPDRS Part III subscore: 0 – 60
    - MDS-UPDRS Part II subscale: 0 – 52
    - MDS-UPDRS Part I subscale: 0 – 52

At the end of the imputation process, **B x M** complete datasets will be available.

For all imputation models, sex (reference = Male), age at informed consent signing (agegrp4, reference = 60-69), and geographic region (reference = Europe) will be included as baseline demographics. No missing data is expected for baseline demographics. MI models will only allow continuous variables. Therefore, categorical variables will be re-coded as indicator variables as follows:

Baseline demographics (included in all analyses)

- For sex: Male = 0, Female = 1
- For age at informed consent: 60-69 = 0, 40-49 = 1, 50-59 = 2, 70+ = 3
- For geographic region: Europe = 0, North America = 1

Baseline/screening data for time varying covariates MDS-UPDRS Part II and DaT-SPECT Striatum SBR (not mean-centered, constrained to be positive) will also be included as baseline covariates. Where missing data are present, a MI-MCMC approach assuming MAR – as for intermittent missing data – will be used. For all variables, imputed datapoint outside the expected range for the variable will be rest to the appropriate lower/upper limit of the range.

Exploratory sensitivity analyses may be run with using additional covariates for the imputation step (details to be included in the exploratory analysis plan).

6.5.5 ST de-mediation via g-estimation

A g-estimation approach will be implemented on each of the **B x M** imputed complete dataset, to de-mediate the MDS-UPDRS data from the confounding effect of the rescue medication (i.e. ST, Levodopa mainly) on the observed effect of the active randomized treatment (Minzasolmin/UCB0599) to obtain the direct effect of this active treatment.
The basic idea of apply a g-estimation algorithm is to "peel off" the effect of ST on all future MDS-UPDRS observations.

Complete post-MI-MCMC MDS-UPDRS Part I-III sum score, Part III subscale and Part III ePD subscore outcome data will be used in the g-estimation models, both as the dependent variable and as the respective lagged outcomes, i.e. time-varying covariate.

Note that to be conservative, RBI data from participants who discontinued the study for reasons related to randomized treatment will not be ST de-mediated.

In the g-estimation approach, sex (reference = Male), age at formed consent signing (agegrp4, reference = 60-69), and geographic region (reference = Europe) will be included as baseline demographic covariates and carried over time. Baseline/screening data for MDS-UPDRS Part II and DaT-SPECT Striatum SBR (not mean-centered) will also be included as time-varying covariates. Where post-baseline data for these time-varying covariates are missing, a last observation carried forward (LOCF) approach will be applied to impute missing data, both in the case where the variable is not measured at a visit by design, or in the case where it is truly missing (assumes MAR).

Exploratory sensitivity analyses may be run where additional time-varying covariates are included in the g-estimation model as lagged predictors (details to be included in the exploratory analysis plan).

- Prepare dataset for g-estimation 2-step modeling:
  - Set time t to start at 1 up to T (where T = 10 visits, Baseline, Month 2 … Month 18)
  - For each **B x M** imputed dataset (including any baseline/post-baseline variables as required), within each participant, lag the following variables over each Visit V_t using SAS functions Lag1, Lag2 to Lag(T-1), i.e.:
    - S_t , whether a participant initiates ST between V_t and V_{t+1}, set to 1 if yes, set to 0 otherwise (generates lag1S, lag2S …); once a participant has initiated ST, it is assumed they stay on ST for the remaining of the trial;
    - Y_t , the participant's outcome measured at V_t (generates lag1Y, lag2Y …)
    - $L_t(s)$, the participant's time-varying variables measured at V_t (generates e.g. lag1updrs2, lag2updrs2 …).
  - For each **B x M** imputed dataset, carry forward participant's demographics L_1 measured at screening/baseline across all visits (generates e.g. lagSex …)
  - Exclude baseline rows (t = 1) from dataset.

- Step 1: Propensity model - For each **B x M** dataset, estimate the participant's propensity score (ps), i.e. the probability of initiating ST at any post-baseline visit based on outcome using a logistic regression model.
  - Exclude rows corresponding to $V_{t=T}$ (i.e. Month 18), where S_t is not applicable
  - Retain rows where S_{t-1} (lag1S) = 0 (i.e. participants did not initiate ST up to and including V_t)
  - Run a logistic regression of S_t on Y_t, adjusting for:
    - V_t, R (Randomized treatment arm) - Class variables as Fixed effects:
    - Sex, Age group, and Geographic region - Class variables as demographics covariates

  - MDS-UPDRS Part II, DaT-SPECT striatum SBR - Continuous variables as baseline covariates
- Use model to predict the ps_t for each participant at each V_t.
  - Where S_{t-1} (lag1S) = 1 or S_t not applicable, then set $ps_t = 1$;
- Lag the ps_t variable within each participant, over each V_t as described in the data preparation section above (lag1PS to lagTPS).

- Step 2: linear structural mean model (LSMM) - For each **B x M** dataset, estimate the lag effects ψ_g of initiating ST between V_t and V_{t+1} on Y_{t+g} for all possible lags g from 1 to G (T-1), assuming ψ_g is time-invariant, by implementing a series of linear regression models and de-mediate the overall lag effects of ST on outcome:

To estimate ψ_1 (example for lag g = 1),

- Retain rows Y_{t+1} where S_{t-1} (lag2S) = 0 for t = 1, …, T-1
- Run a linear regression of Y_{t+1} on 1-lag outcome Y_t (lag1Y), observed ST initiation S_t (lag1S), and propensity score for ST initiation ps_t (lag1PS), adjusting for:
  - R (Randomized treatment arm) - Class variables as Fixed effects (Visit not included as fixed effect):
  - Sex, Age group, and Geographic region - Class variables as demographics covariates
  - MDS-UPDRS Part II, DaT-SPECT striatum SBR - Continuous variables as baseline covariates
- Set ψ_1 estimate equal to the linear coefficient of S_t (lag1S)

To estimate ψ_g for g = 2, …, T-1 (generalization to all possible lags)

- Retain rows $\tilde{Y}_{t+g}$ where S_{t-1} (lag(***g+1***)S) = 0 for t = 1, …, T-g
- Run a linear regression of $\tilde{Y}_{t+g}$ on g-lag outcome Y_t (lag***g***Y), observed ST initiation S_t (lag***g***S), and propensity score for ST initiation ps_t (lag***g***PS), adjusting as above, for t = 1, …, T-g in a single step
- Set ψ_g estimates equal to the linear coefficients of S_t (lag***g***S)

- Step 3: obtain de-mediated outcome data points
  - Peel off lag effect ψ_1 from outcome Y_{t+1} to obtain the *1-lag de-mediated* outcome:
    $\tilde{Y}_{t+1} = Y_{t+1} - \psi_1 \times \text{lag1S}$ for t = 1, …, T-1
  - Update $\tilde{Y}_{t+g}$ to sequentially remove all lag effects:
    $\tilde{Y}_{t+g} = \tilde{Y}_{t+g} - \psi_g \times \text{lag}\boldsymbol{g}\text{S}$ for t = 1, …, T-g

  In practice, $\hat{Y}_{t+g}$ will only be updated where lag***g***S = 1

6.5.6 End-of-trial Analysis Models

Repeated measures model

A mixed effect model (MEM) for longitudinal analysis of covariance (ANCOVA)/repeated measures will be applied to each **B x M** bootstrapped/imputed/ST de-mediated dataset to obtain **B x M** mean differences (i.e., $\theta_{b,m}$, see section 6.5.7) at Month 12 and Month 18 - as required per MDS-UPDRS endpoint – using all data recorded up to Month 12/Month18, respectively.

Treatment, time (visit, categorical) and treatment by time (interaction term between treatment and visit) will be fitted as fixed effects, gender, age at Baseline and respective Baseline MDS-UPDRS data will be fitted as covariates. A heterogenous auto-regressive (ARH1) variance-covariance matrix will be fitted to account for the repeated measures within subject (SAS PROC MIXED with REPEATED statement). Kenward and Roger adjustment will be used to obtain degrees of freedom.
For further details, refer to Section 5.3.1.4.1.

Slope model
A LMEM (random coefficients model) for longitudinal data will be applied to each of the B x M bootstrapped/imputed/ST de-mediated dataset to obtain **B x M** mean differences (i.e., $\theta_{b,m}$, see section 6.5.7) in slope over 12 and/or 18 months - as required per MDS-UPDRS endpoint – using all data recorded up to Month 12/Month18, respectively.
MDS-UPDRS will be the dependent variable, Baseline MDS-UPDRS data will be part of the dependent variable rather than a covariate in the model. Time since Baseline in months - as a continuous variable - and the treatment by time interaction (fixed slope) will be fitted as fixed effects, gender and age at Baseline will be fitted as covariates. A fixed effect for treatment will NOT be included in this model. To adjust for the correlation between repeated observations from each participant, a participant-specific random intercept and a participant-specific random slope for time will be fitted as random effects (SAS PROC MIXED with RANDOM statement with option type = UN Kenward and Roger adjustment will be used to obtain degrees of freedom.
For further details, refer to Section 5.3.1.4.1.

6.5.7 Results Summaries

The results from the MEMs performed on each of the **B x M** bootstrapped/imputed/ST de-mediated dataset will then be combined for overall inference using the Boot MI percentile approach (section 3.2.1 in Bartlett & Hughes 2020).

Obtain dataset-specific mean point estimates θ_b of interest for each of the B dataset over the M imputations,

$$\bar{\theta}_b = \frac{1}{M}\sum_{m} \hat{\theta}_{b,m}$$

and overall mean point estimate θ over the B datasets:

$$\bar{\theta} \quad = B^{-1}\sum_{b=1}^{B} \bar{\theta}_b, \text{ where } \bar{\theta}_b = M^{-1}\sum_{m=1}^{M} \hat{\theta}_{b,m}$$

Where $\theta_{b,m}$ represents the point estimate of the parameter of interest for a given endpoint obtained by applying the analysis model to any of the B x M datasets.

$\bar{\theta}$ corresponds to the average of the B x M parameter estimates (θ_b), and its 95% CI is derived by obtaining the 2.5% and 97.5% bootstrap percentile intervals of θ_b. Its 80% CI is derived by obtaining the 10% and 90% bootstrap percentile intervals of θ_b.

Derivation of 95% CIs and 80% CIs will use the same approach regardless of the parameter of interest, i.e. mean at single timepoint, mean slope, differences in mean at single timepoint or difference in mean slope.

Descriptive statistics

Summary tables presenting the observed mean and mean change from Baseline in ST de-mediated MDS-UPDRS Part I-III sum score, Part III subscale and Part III ePD subscore data will be produced by treatment group and visit.

Plots of observed mean and mean change from baseline over time will be produced by treatment group. These plots will include error bars (percentile intervals) and all treatment groups will be overlaid on the same plot.

Repeated measures model

A summary table presenting the estimated means and treatment effects of interest, i.e. difference in target population means at Month 12/Month 18 (based on the treatment by time interactions) and corresponding 95% CIs/80% CIs will be produced:

- UCB0599 high dose (360mg/day) and placebo,
- UCB0599 low dose (180mg/day) and placebo.

The % change relative placebo will also be included (see section 5.3).

A plot displaying the estimated ST de-mediated means and 95% CIs/80% CIs for each treatment group at each time point up to Month 12/Month 18, as well as differences in means and corresponding 95% CIs/80% CIs for Month 12/18 and will be produced.

For further details, refer to section Section 5.3.1.4.1.

Slope model

A summary table presenting the estimated means and treatment effects of interest, i.e. difference in target population mean slope of progression over 12/18 months (based on the treatment by time interactions) and corresponding 95% CIs/80% CIs will be produced for:

- UCB0599 high dose (360mg/day) and placebo,
- UCB0599 low dose (180mg/day) and placebo.

A plot displaying the adjusted ST de-mediated mean slopes and 95% CIs/80% CIs for each treatment group from this model as well as differences in mean slope and corresponding 95% CIs/80% CIs will be produced.

For further details, refer to Section 5.3.1.4.1.

6.6 Appendix 6: Key inclusion/exclusion criteria for the FAS definition

Participants who did not meet the key inclusion criteria stated below will be excluded from the FAS:

Protocol Criterion 1a: Study participant's age at baseline must be >39 and <76 years, i.e. Study participant must be 40 to 75 years of age inclusive, at the time of signing the informed consent.

Protocol Criterion 2a: Study participant has PD, with a diagnosis made by a neurologist according to the 2015 Movement Disorder Society criteria within 2 years of Baseline Visit (including diagnosis during Screening), at the time of signing the informed consent.

In practice, this be implemented by excluding participants with a disease duration greater than 27 months (3 months over the protocol-defined threshold).

Protocol Criterion 4b: A Screening DaT-SPECT, or a historical DaT-SPECT within 3 months of the Screening Visit (V1) that has been qualified by the central reader, shows evidence of dopamine transporter deficit per study requirements (refer to Section 4.2 of the protocol) and as determined by a central reader.

Protocol Criterion 6a: Study participant is in the ≤2.5 modified Hoehn and Yahr stage at Screening.

Criterion 7a. Study participant has never taken medications for the treatment of motor symptoms of PD and is not expected to require starting ST with a high likelihood in the next 6 months as far as clinical judgement allows.

7 REFERENCES

Ahn C, Heo M, and Zhang, S. (2014.) Chapter 5 Section 5.4.2.2 Unbalanced allocations. In: Sample size calculations for clustered and longitudinal outcomes in clinical research. New York: Chapman and Hall. 10.1201/b17822.

Bartlett JW, Hughes RA. Bootstrap inference for multiple imputation under uncongeniality and misspecification. Stat Methods Med Res. 2020 Dec;29(12):3533-3546. doi: 10.1177/0962280220932189. Epub 2020 Jun 30. PMID: 32605503; PMCID: PMC7682506.

Brown H and Prescott R. (2014.) Applied Mixed Models in Medicine. (3 ed.) (Statistics in Practice) John Wiley & Sons Inc. 10.1002/9781118778210.

CPMP/ICH/363/96. Note for guidance on Statistical Principles for Clinical Trials (EMA). September 1998

Cummings J. Disease modification and neuroprotection in neurodegenerative disorders. Transl Neurodegener. 2017;25: Article number 25.

EuroQol Group 2017, EQ-5D-5L About, viewed 9 November 2020, < https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/>

Floden L, Bell ML. Imputation strategies when a continuous outcome is to be dichotomized for responder analysis: a simulation study. BMC Med Res Methodol. 2019 Jul 23;19(1):161.

Horváth K, Aschermann Z, Ács P, Deli G, Janszky J, Komoly S, Balázs É, Takács K, Karádi K, Kovács N. Minimal clinically important difference on the Motor Examination part of MDS-UPDRS. Parkinsonism Relat Disord. 2015;21(12):1421-1426.Ikeda K, Ebina J, Kawabe K,

Iwasaki Y. Dopamine Transporter Imaging in Parkinson Disease: Progressive Changes and Therapeutic Modification after Anti-parkinsonian Medications. Intern Med. 2019;58(12):1665-72.

Lasch, F., Guizzaro, L., Pétavy, F., & Gallo, C. (2023). A Simulation Study on the Estimation of the Effect in the Hypothetical Scenario of No Use of Symptomatic Treatment in Trials for Disease-Modifying Agents for Alzheimer's Disease. *Statistics in Biopharmaceutical Research*, *15* (2), 386-399.

Lingling Li. SAS® V9.4 MNAR statement for multiple imputations for missing not at random in longitudinal clinical trials. PharmaSUG 2019 - Paper ST-103.

Liu-Seifert H, Andersen SW, Lipkovich I, Holdridge KC, Siemers E. A novel approach to delayed-start analyses for demonstrating disease-modifying effects in Alzheimer's disease. PLoS One. 2015;10(3):e0119632.

Makkos A, Kovács M, Aschermann Z, et al. Are the MDS-UPDRS-Based Composite Scores Clinically Applicable?. *Mov Disord*. 2018;33(5):835-839.

Mehrotra DV, Liu F, Permutt T. Missing data in clinical trials: control-based mean imputation and sensitivity analysis. Pharm Stat. 2017;16(5):378-92.

Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief Screening tool for mild cognitive impairment. J Am Geriatr Soc. 2005;53(4):695-9.

Royston, P, Parmar, MK. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. BMC Med Res Methodol. 2013;13:152 .

Parashos SA, Swearingen CJ, Biglan KM, Bodis-Wollner I, Liang GS, Ross GW, et al. Determinants of the timing of symptomatic treatment in early Parkinson disease: The National Institutes of Health Exploratory Trials in Parkinson Disease (NET-PD) Experience. Arch Neurol. 2009; 66(9):1099-104.

Parkinson's Progression Markers Initiative 2018, viewed 24 November 2020, <http://www.ppmi-info.org/>

Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168:1266-77.

Simuni T, Long JD, Caspell-Garcia C, Coffey CS, Lasch S, Tanner CM, et al. Predictors of time to initiation of symptomatic therapy in early Parkinson's disease. Ann Clin Transl Neurol. 2016;3(7):482-94.

Tosin, Michelle H.S. et al. Tracking Emergence of New Motor and Non-Motor Symptoms Using the MDS-UPDRS: A Novel Outcome Measure for Early Parkinson's Disease? 1 Jan. 2022;1345–1351.

Twisk, J. (2013). Chapter 5: The modeling of time. In: Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide (2nd ed.). Cambridge: Cambridge University Press.

Twisk J, Bosman L, Hoekstra T, Rijnhart J, Welten M, Heymans M. Different ways to estimate treatment effects in randomised controlled trials. Contemp Clin Trials Commun. 2018;10:80-5.

Van Lancker K, Dukes O, Vansteelandt S. Principled selection of baseline covariates to account for censoring in randomized trials with a survival endpoint. Stat Med. 2021 Aug 15;40(18):4108-4121. doi: 10.1002/sim.9017. Epub 2021 May 12. PMID: 33978249.

Vansteelandt, Stijn and Marshall M. Joffe. "Structural nested models and G-estimation: the partially realized promise." Statistical Science 29 (2014): 707-731.

Wasserstein RL, Lazar NA. The ASA Statement on *p*-Values: Context, Process, and Purpose, The American Statistician. 2016; 70(2):129-33.

White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials. 2012;9(4):396-407.

Approval Signatures

Name: PD0053-sap-amend-4

Version: 1.0

Document Number: CLIN-000266646

Title: PD0053 SAP Amendment 4

Approved Date: 21 Oct 2024

Document Approvals	
Approval Verdict: Approved	Name: [redacted] Capacity: Subject Matter Expert Date of Signature: 21-Oct-2024 15:47:16 GMT+0000
Approval Verdict: Approved	Name: [redacted] Capacity: Clinical Date of Signature: 21-Oct-2024 18:09:11 GMT+0000