Medtronic Engineering the extraordinary

# Statistical Analysis Plan Cover Page

## **A**daptive **D**BS **A**lgorithm for **P**ersonalized **T**herapy in **P**arkinson's **D**isease (ADAPT-PD) Trial

Clinicaltrials.gov ID: NCT04547712

Document Date: 19 October 2023

| Medtronic<br>**Statistical Analysis Plan** | |
|---|---|
| Clinical Investigation Plan Title | **A**daptive **D**BS **A**lgorithm for **P**ersonalized **T**herapy in **P**arkinson's **D**isease (ADAPT-PD) Trial |
| Clinical Investigation Plan Identifier | MDT19001<br>EUDAMED identifier will be provided under a separate cover when available |
| Clinical Investigation Plan Version | Version 5.0 and subsequent versions |
| Statistical Analysis Plan Version Date | October 19, 2023 |
| Sponsor/Local Sponsor | Medtronic<br>Medtronic Neuromodulation<br>7000 Central Ave. NE<br>Minneapolis, MN 55432 USA<br><br>Medtronic International Trading Sàrl<br>Route du Molliau 31<br>1131 Tolochenaz, Switzerland<br><br>EU Legal Representative<br>Medtronic Bakken Research<br>Center Endepolsdomein 5<br>6229 GW Maastricht The Netherlands<br><br>Medtronic Canada ULC<br>99 Hereford Street<br>Brampton, Ontario L6Y 0R3 Canada<br>+1-905-460-3800 |

# Table of Contents

# 1. Version History

| Version | Summary of Changes | Author(s)/Title |
|---|---|---|
| **1.0** | • Not Applicable, New Document | [redacted] |
| **2.0** | Aligned with Version 3.0 of CIP, including:<br>• Added the Directional Stimulation cohort<br>• Added TEED calculations for subjects implanted with the SenSight system<br>• Added poolability analyses of a lead type effect (legacy Models 3387 and 3389 compared to SenSight Models B33005 and B33015) for the primary and secondary objectives and safety assessment<br>• Added supportive analyses of the Directional Stimulation Cohort for the primary and secondary objectives and safety assessment<br>• Updated the Patient Preference and Patient Satisfaction analyses and screenshots<br><br>Other updates include:<br>• Specified that PD Home Diary complete diary periods collected anytime during cDBS will be used for analysis<br>• Added examples of coding that could be used for multiple imputation analyses, tests of a period or carryover effect, and poolability tests of the primary and secondary objectives and safety assessment<br>• Defined the safety evaluation populations | [redacted] |
| **3.0** | • Updated the document using the current SAP template<br>• Added SAP version date to cover page per ClinicalTrials.gov requirements<br>• Editorial updates for clarifications | [redacted] |
| **4.0** | Aligned with Version 5.0 of CIP, including:<br>• Allowed for impedance values from aDBS Setup Visit to be used for analysis in case they were not collected at cDBS Baseline or aDBS Evaluation Phase<br>• Added a subgroup analysis of outcomes during Evaluation Phase by the mode selected for the Long-term Follow-up | [redacted] |

## 2. List of Abbreviations and Definitions of Terms

| Abbreviation | Definition |
|---|---|
| AC | All-Consented |
| AE | Adverse Event |
| aDBS | Adaptive Deep Brain Stimulation |
| AR | All-Randomized |
| AT | As-Treated |
| Best condition | The time period when PD subjects are receiving relief from their Parkinson's disease symptoms, also known as "on" time or "on" period |
| BKS | Bradykinesia Score |
| cDBS | Continuous Deep Brain Stimulation |
| CC | Complete-Case |
| CIP | Clinical Investigation Plan |
| CONSORT | Consolidated Standards of Reporting Trials |
| CRF | Case Report Form |
| DBS | Deep Brain Stimulation |
| DS | Directional Stimulation Cohort comprised of study subjects with the SenSight™ Directional Lead programmed to directional stimulation |
| DKS | Dyskinesia Score |
| EQ-5D-5L | Five-dimensional five-level generic measure designed to measure, compare and value health status across disease areas; developed by the EuroQol Group. |
| Event markers | Used to capture and record patient triggered events and provide additional information for analysis around patients outside the clinic. |
| FAS | Full Analysis Set |
| FCS | Fully Conditional Specification |
| FDS | Fluctuation and Dyskinesia Score |
| GIC | Global Impression of Change |
| GPi | Globus Pallidus Internus |
| INS | Implantable Neurostimulator |
| ITT | Intention-to-Treat |
| JSON | JavaScript Object Notation, a data file format |
| LFP | Local Field Potential |
| MDS-UPDRS | Movement Disorder Society Unified Parkinson's Disease Rating Scale |
| MI | Multiple Imputation |

| Abbreviation | Definition |
|---|---|
| PD | Parkinson's Disease |
| PDQ-39 | The 39-item Parkinson's Disease Questionnaire |
| PDQ-39 SI | PDQ-39 summary index |
| PDSS-2 | Parkinson's Disease Sleep Scale 2 |
| Primary Cohort | Cohort comprised of study subjects implanted with legacy lead models 3387 or 3389, or with the SenSight™ Directional Lead programmed to ring mode stimulation |
| SAP | Statistical Analysis Plan |
| STN | Subthalamic Nucleus |
| VAS | Visual Analog Scale |
| VHI | Voice Handicap Index |
| Worst condition | The time period when PD subjects are not receiving relief from their Parkinson's disease symptoms, also known as "off" time or "off" period. |

# 3. Introduction

## 3.1 Introduction

The ADAPT-PD trial is a prospective, multicenter, single-blind, randomized crossover clinical investigation and is designed to demonstrate the safety and effectiveness of adaptive deep brain stimulation (DBS) for Parkinson's disease. It is composed of a primary trial cohort (Primary cohort) and a cohort of subjects implanted with the SenSight Directional Lead system and programmed to directional stimulation (Directional Stimulation Cohort). The study duration is expected to be approximately 5 years. It is estimated that approximately 16 months will be needed to enroll 70-100 subjects in order to obtain 40 subjects for whom at least 1 adaptive DBS (aDBS) mode is acceptable and who will enter the aDBS Evaluation Phase of the Primary cohort and to enroll approximately 15 subjects in order to obtain 9 subjects for whom at least 1 aDBS mode is acceptable and a minimum of 8 who complete the aDBS Evaluations Phase in the Directional Stimulation Cohort.

## 3.2 Purpose

The purpose of this statistical analysis plan (SAP) is to document the analyses for the primary objective report used for regulatory submission and for the final study report. Revisions to the SAP may be required if the protocol changes or updates to the analysis are needed. Annual progress reports may contain a subset of the analyses described in this SAP (e.g., demographics and baseline characteristics, safety evaluation, deviations, early discontinuations). The study objectives are taken directly from the

Clinical Investigation Plan (CIP). The SAP will further define safety and effectiveness analyses, starting in section 7.9.

## 3.3 Documents used to create SAP

Documents used to create the SAP include:

- ADAPT-PD Clinical Investigational Plan

# 4. Study Objectives

## 4.1 Primary objective

To demonstrate that the proportion of aDBS subjects with "On" time without troublesome dyskinesia during the Evaluation Phase exceeds a performance goal of 50%.

## 4.2 Secondary objective

To demonstrate decreased stimulation energy use during the aDBS Evaluation Phase as compared with continuous DBS (cDBS).

## 4.3 Safety analyses

To characterize:

- Stimulation-related adverse events during the aDBS Evaluation and the cDBS Baseline Phases.
- Serious adverse events, adverse events and device deficiencies throughout the study.

## 4.4 Additional objectives

To characterize aDBS during the Evaluation Phase as compared to cDBS (Primary Cohort):

- Voice Handicap Index (VHI): "best" and "worst" conditions
- Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Part III subscore (On stim/On med) of tremor, rigidity, and bradykinesia questions where a side can be determined: 3.3 b, c, d, e (rigidity), 3.4 a, b (finger taps/bradykinesia), 3.15 a, b (postural tremor of hands), 3.16 a, b (kinetic tremor of hands), and 3.17 a, b, c, d (rest tremor amplitude)
- MDS-UPDRS III (On stim/On med)
- MDS-UPDRS II: "best" and "worst" conditions
- MDS-UPDRS III speech question (On stim/On med) q 3.1
- MDS-UPDRS IV
- European Quality of Life – 5 Dimensions, version 5L (EQ-5D-5L)
- Parkinson's Disease Sleep Scale 2 (PDSS-2)
- Data collected from wearable (for example: % of wear time tremor was detected, Dyskinesia Score [DKS], Bradykinesia Score [BKS], Fluctuation and Dyskinesia Score [FDS])

- Parkinson's Disease Questionnaire (PDQ)-39 Summary Index (SI) and subscores: mobility, activities of daily living, emotional well-being, stigma, social support, cognition, communication, bodily discomfort
- PDQ-39 speech questions 34 and 35
- % of subjects with at least 25% increase in projected battery longevity as compared with cDBS
- Parkinson's Disease (PD) Home Diary: "On" time without troublesome dyskinesia, "On" time without dyskinesia, "On" time with non-troublesome dyskinesia, "On" time with troublesome dyskinesia, "Off" time, and Asleep time.
- Patient preference of aDBS vs cDBS (as assessed by a Medtronic-developed patient preference questionnaire)
- Patient satisfaction with aDBS (as assessed by a Medtronic-developed patient satisfaction questionnaire)
- Device data from JavaScript Object Notation (JSON) file, including BrainSense data

Subgroup analyses (from the entire study population):

- For those with dyskinesia at least 25% of time: Wearable data
- For those with speech side effects: VHI
- For those with mixed programming (aDBS and cDBS): MDS-UPDRS III by side
- For those who use event markers in long-term follow-up: characterize subject population and satisfaction with use of event markers
- For those who finished the aDBS Evaluation Phase: one aDBS mode will be selected per subject based on subject preference or programming availability at the end of the aDBS Evaluation Phase. The aDBS Evaluation Phase results may be characterized based on the selected mode.

# 5. Investigation Plan

This is a prospective, single-blind, randomized crossover, multi-center study of aDBS in subjects with Parkinson's disease. It is composed of a primary trial cohort (Primary cohort) and a cohort of subjects implanted with the SenSight system and programmed to directional stimulation (Directional Stimulation Cohort).

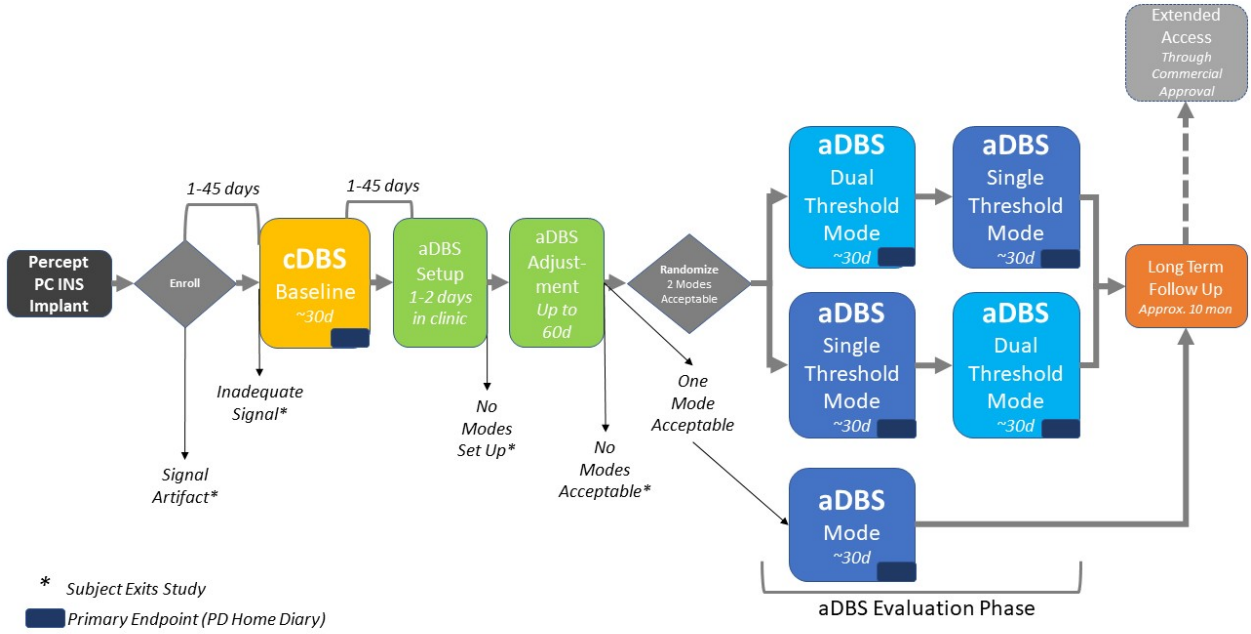
The study is expected to be conducted at approximately 12 centers located in the US, Europe and Canada. It is estimated that approximately 70-100 subjects implanted with a commercially available Medtronic DBS system with Percept PC implantable neurostimulator (INS) will be enrolled in the Primary Cohort to obtain 40 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase. Accounting for a 10% dropout during the aDBS Evaluation Phase, a minimum of 36 subjects (at least 8 per brain target) will complete the aDBS Evaluation Phase. In addition, approximately 15 subjects implanted with a commercially available Medtronic DBS system with Percept PC INS and SenSight system will be enrolled in the Directional Stimulation Cohort to obtain 9 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase and a minimum of 8 subjects who complete the aDBS Evaluation Phase.

Study visits and/or phases include (see Figure 5-1):

- Enrollment Visit: Consent, Screening
- cDBS Baseline Phase: Local Field Potential (LFP) Screening and Baseline cDBS visits
- aDBS Setup and Adjustment Phase: aDBS Setup visit and additional optional visits during an aDBS Adjustment period
- aDBS Evaluation Phase: One-month treatment periods in each acceptable aDBS mode with aDBS evaluation visits at the end of each period. Randomized crossover to the two investigational treatments in subjects for whom both aDBS modes were acceptable and a single treatment period in those subjects for whom only one aDBS mode was acceptable.
- Long-Term Follow Up Phase: Four visits in preferred aDBS mode for approximately 10 months
- Extended Access Phase: Additional visits in the preferred aDBS mode every 6 months through commercial approval of aDBS

The design of the Directional Stimulation Cohort is the same as the main ADAPT-PD trial. Subjects enrolled in the Directional Stimulation Cohort will follow the same visit schedule.

**Figure 5-1. Study Schematic**




# 6. Determination of Sample Size

Since the device implant will occur outside of the study and the study requires participants to be stable on cDBS prior to enrollment, the standard deviation of the change from cDBS to aDBS in "On" time

without troublesome dyskinesia will be used to establish a threshold for assessing the performance goal. The threshold will be set as 1 standard deviation. If the difference between aDBS and cDBS (aDBS-cDBS) for a subject is at least -1 standard deviation, the threshold will be met for the subject. Using this threshold, the proportion of subjects who exceed the threshold will be computed.

The sample size was estimated using a binomial distribution for a one-sided exact test (alpha=0.0125) for a proportion (PASS 2011 module Tests for One Proportion [Differences]) to compare to a performance goal of 50%. Assuming the alternative hypothesis of 85% of subjects exceeding the threshold, a minimum of 36 subjects achieves at least 90% power to reject a performance goal of 50%.

Simulations have shown that this sample size is sufficient to cover possible deviations from normality such as a uniform distribution. The threshold criterion will remain the same as the standard deviation is robust to moderate departures from normality.

Sample size in the Primary Cohort may be increased to a maximum of 55 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase to obtain at least 8 subjects per aDBS mode and 8 subjects per target site (Subthalamic Nucleus [STN]/ Globus Pallidus [GPi]). Criteria for increasing sample size is described in section 7.2.1.

For the safety analyses, a minimum of 30 subjects followed for approximately 30 days with aDBS out-of-clinic in the subjects' real-world environments provides at least 95% probability that all events with a true event rate of at least 10% would be reported at least once.

It is estimated that approximately 70-100 subjects implanted with a commercially available Medtronic DBS system with Percept PC INS will be enrolled in the study to obtain 40 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase. Accounting for a 10% dropout during the aDBS Evaluation Phase, a minimum of 36 subjects (at least 8 per brain target site (STN/GPi) and 8 per aDBS mode) will complete the aDBS Evaluation phase.

In addition, approximately 15 subjects implanted with a commercially available Medtronic DBS system with Percept PC INS and SenSight system will be enrolled in the Directional Stimulation Cohort to obtain 9 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase. Accounting for dropout during the aDBS Evaluation Phase, a minimum of 8 subjects will complete the aDBS Evaluation Phase.

# 7. Statistical Methods

## 7.1 Study Subjects

### 7.1.1 Disposition of Subjects

Subject disposition will be illustrated in a Consolidated Standards of Reporting Trials (CONSORT) diagram and attrition will be identified and summarized. The number of subjects screened, randomized, and receiving therapy will be summarized.

### 7.1.2 Clinical Investigation Plan (CIP) Deviations

CIP deviations will be listed and the number of deviations will be summarized by type and overall.

### 7.1.3 Analysis Sets

All-Consented (AC): includes all subjects who properly signed the study-specific informed consent (IC). This patient set will be used to summarize patient disposition.

Full Analysis Set (FAS): includes all Primary Cohort subjects, using the intention-to-treat (ITT) principle, who initiate the aDBS Evaluation Phase using the randomized treatment assignment for each subject that was randomized and the programmable treatment assignment for those subjects that may only be configured to one aDBS mode (dual or single threshold).

As-Treated (AT): includes all Primary Cohort subjects who are treated with cDBS and/or aDBS during the cDBS Baseline Evaluation Phase and/or the aDBS Evaluation Phase, and Long-term Follow-up Phase and uses the observed treatment used for each subject in each phase.

All-Randomized (AR): includes all Primary Cohort subjects who were randomized in the aDBS Evaluation Phase, excludes those that were not configured to both modes.

Complete-Case (CC): FAS without imputation that includes all Primary Cohort subjects who have available measures for the respective analysis and are analyzed in their randomization group for those randomized or the programmable treatment for those configured to one aDBS mode.

Directional Stimulation As-Treated (DSAT): includes all Directional Stimulation Cohort subjects who have SenSight leads in a directional stimulation configuration who are treated with continuous Deep Brain Stimulation (cDBS) and/or adaptive Deep Brain Stimulation (aDBS) during the cDBS Baseline Evaluation Phase and/or the aDBS Evaluation Phase, and Long-term Follow-up Phase and uses the observed treatment used for each subject in each phase.

Directional Stimulation Complete-Case (DSCC): includes all Directional Stimulation Cohort subjects who completed the aDBS Evaluation Phase.

## 7.2 General Methodology

The ADAPT-PD Trial will be considered successful when the primary objective of the study is met for one or both aDBS algorithms independently or the algorithms combined.

The main analysis of the study objectives will use the FAS which includes all Primary Cohort subjects, using the ITT principle, who initiate the aDBS Evaluation Phase using the randomized treatment assignment for each subject that was randomized and the programmable treatment assignment for those subjects that may only be configured to one aDBS mode (dual or single threshold).

The main analysis will include data from all contributing geographies and centers. A poolability analysis for a center effect will be performed as described in Section 7.3.

The subjects' baseline status will be summarized with standard descriptive summary statistics, including counts and percentages for categorical variables, and mean, standard deviation for continuous variables. Safety data will be summarized as specified in Section 7.10.

### 7.2.1 Pooling of aDBS mode

Most planned analyses for the Primary Cohort are to be done separately for each mode. Summary statistics will be computed to summarize the percentage who were programmed to each mode and subject characteristics of each subgroup. If fewer than 24 subjects are programmed to one of the aDBS modes because aDBS is unacceptable as compared to cDBS (using the aDBS Global Impression of Change [GIC]), the evaluation of aDBS effectiveness will use data from the first period of the aDBS Evaluation Phase from all subjects. If the modes are considered poolable, the modes would be combined and the alpha level of the test would be alpha=0.025 (one-sided). To test for a treatment difference in the first period by aDBS mode group, subjects will be categorized based on the modes they are programmable to: aDBS – Dual Threshold only; aDBS – Single Threshold only; and aDBS – both modes. The Freeman and Halton's extension of Fisher's exact test will be computed testing for differences between aDBS modes in the proportion of subjects exceeding threshold for the primary outcome using data only from the first period. [redacted]

The same categorization of mode will be used and tested using F-tests from linear regression models for the secondary outcome measure. [redacted]

P-values >0.1 for the tests of association with aDBS mode group will be deemed as supporting the assumption of no aDBS mode group effect and the data will be considered poolable across aDBS mode groups. If the aDBS mode group factor approaches statistical significance (defined as ≤0.1),[1] the data will not be considered poolable. A minimum of 8 subjects per aDBS mode will be required for the pooled analysis using the first period outcome data.

If data are pooled, missing data observed in the cDBS Baseline Phase and aDBS Evaluation Period 1 will be imputed as described in 7.4.

## 7.3 Center Pooling

All investigators in the proposed study will conduct the study according to a common protocol and utilize the same Case Report Forms (CRFs) to collect study data. Throughout the trial, efforts will be made to ensure consistency among investigative sites in selection of patients and conduct of the study procedures. In addition, site study personnel training will be conducted prior to initiation of the study at each site and periodic monitoring will be conducted by Medtronic to ensure compliance with protocol requirements.

Data will be pooled across centers for all analyses. Each study center will stop enrollment in the Primary Cohort after they have identified a maximum of 8 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase (~25% of the total planned number of subjects). For the Directional Stimulation Cohort, each study center will stop enrollment after they have identified a maximum of 5 subjects for whom at least 1 aDBS mode is acceptable and who enter the aDBS Evaluation Phase. The actual number of subjects at a study center may be slightly larger due to the number of subjects that are enrolled prior to cessation of enrollment. This restriction is intended to reduce the possibility that the results from one study center will be overly influential in the overall study results.

Tests for a treatment difference in the Primary Cohort among centers are described in the supporting analyses of the objective sections below. When testing for differences between centers, centers with 5 or more FAS subjects will each be tested as separate centers. Those with less than 5 FAS subjects will be combined into a “pooled” pseudo-center to minimize the impact of small samples on the analysis. If the

[1] Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res. 1993; 2(2):121-45.

pseudo-center contains more than 50% of the subjects, the centers will be combined into more than one pseudo-center. If more than one pseudo-center is needed, the centers will be randomly ordered and divided as near the midpoint as possible. If ambiguity between assigning a center to the first or second pseudo-center exits, the center will be assigned to the first pseudo-center.

## 7.4 Handling of Missing, Unused, and Spurious Data and Dropouts

All attempts will be made to minimize missing data. If 5% of the data or fewer are missing for the analysis in the Primary Cohort, no imputation will be used[2]. Otherwise, if missing data are observed, multiple imputation will be used for the analysis of the primary and secondary objectives in the Primary Cohort, as described in the objective sections below. If an aDBS mode is not programmable to a subject, no imputation will occur for those subjects for that aDBS mode. No imputations will be performed for the Directional Stimulation Cohort.

Appropriate confidence limits will be computed to account for the multiple imputations.[3,4] A Complete-Case Analysis Set and As-Treated Analysis Set will be used for sensitivity analyses of the primary and secondary objective analyses in the Primary Cohort.

## 7.5 Adjustments for Multiple Comparisons

In the Primary Cohort, the familywise error rate will be controlled using one-sided tests with an overall alpha-level of 0.025. Each primary hypothesis (one for each aDBS mode) will be assigned a Bonferroni corrected alpha (0.0125) for the respective primary hypothesis (0.025/2). The primary hypothesis will serve as a gatekeeper for the secondary hypothesis within each mode. No adjustments of multiple comparisons are needed for the Directional Stimulation Cohort as no hypothesis testing will be conducted.

## 7.6 Demographic and Other Baseline Characteristics

Demographics and baseline characteristics will be summarized for the AC, FAS, and DSCC analysis sets. This will at least include:

Baseline demographics:

- Age
- Sex
- Ethnicity
- Race

PD history:

[2] Buhi, E, Goodson, P, Neilands T. Out of Sight, Not Out of Mind: Strategies for Handling Missing Data. Am J Health Behav. 2008 Jan-Feb;32(1):83-92, doi: 10.5555/ajhb.2008.32.1.83.
[3] Rubin, D. B. (1976). "Inference and Missing Data." Biometrika 63:581–592.
[4] Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons.

- PD duration
- Dyskinesia duration
- Location of dyskinesias
- Motor fluctuation duration
- Duration of L-DOPA treatment
- Duration of dopamine agonist treatment

## 7.7 Treatment Characteristics

Device data will be summarized as specified in Section 7.9.3.11.

## 7.8 Interim Analyses

No formal interim analysis is planned.

## 7.9 Evaluation of Objectives

### 7.9.1 Primary Objective

To demonstrate that the proportion of aDBS subjects with "On" time without troublesome dyskinesia during the Evaluation Phase exceeds a performance goal of 50%.

#### 7.9.1.1 Hypothesis

Two hypotheses are being tested for the primary objective in order to compare each aDBS mode relative to cDBS at baseline, separately. The alternative hypotheses are:

- $H_{11}$: The proportion of subjects with "On" time without troublesome dyskinesia during aDBS dual threshold mode Evaluation Period exceeding threshold > 50%
- $H_{21}$: The proportion of subjects with "On" time without troublesome dyskinesia during aDBS single threshold mode Evaluation Period exceeding threshold > 50%

#### 7.9.1.2 Endpoint definition

The primary endpoint is the proportion of aDBS subjects with "On" time without troublesome dyskinesia that exceed threshold. This endpoint will combine the diary categories of "On" time without dyskinesia and "On" time with non-troublesome dyskinesia. Since the device implant will occur outside of the study and the study requires participants to be stable on cDBS, the within-subject standard deviation of the change between cDBS and DBS in "On" time without troublesome dyskinesia will be used to establish a threshold for assessing the performance goal. The threshold will be set as 1 standard deviation. If the difference between aDBS and cDBS (aDBS-cDBS) for a subject is at least -1 standard deviation, the threshold will be met for the subject. Using this threshold, the proportion of subjects who exceed the threshold will be computed. The primary endpoint proportion of subjects above threshold will be compared to a performance goal of 50% using a binomial exact test. The lower bound of the 97.5%

confidence limit will also be presented. Data will be imputed as described in 7.4 prior to computing the mean "On" time without troublesome dyskinesia for all periods when using the FAS.

**Rationale for endpoint**

Continuous Deep Brain Stimulation is a proven FDA-approved therapy in the treatment of Parkinson's disease symptoms. This study aims to show that aDBS is a viable optional feature that can provide similar benefit and safety to cDBS for some subjects. The primary endpoint for this study compares aDBS to cDBS directly for each subject to ensure some level of consistent comparability between cDBS and aDBS. The endpoint of PD Home Diary was chosen to determine the threshold of success as it is a measure of the subject response to DBS therapy over time, the aDBS algorithm will adapt the therapy over time, and the PD Home Diary has been used as an endpoint in previous DBS trials.

*Rationale for 1 standard deviation threshold*

A unit of one standard deviation was chosen as it is an acceptable amount of statistical variability in the PD Home Diary endpoint given the fluctuating nature of Parkinson's disease and the percentiles of a normal distribution.

*Rationale for 50% performance goal*

A performance goal of 50% was chosen for this study in order to demonstrate that at least 50% of the subjects are able to achieve the threshold with 95% confidence demonstrating some level of efficacy.

#### 7.9.1.3 Sample size justification

The sample size justification for the primary objective is provided in section 6.

#### 7.9.1.4 Analysis methods

The primary analysis will use the FAS as described in section 7.1.3. The PD Home Diary will be collected for at least 3 consecutive days (24 hour "periods") prior to each visit to evaluate the primary objective. Complete diary periods collected anytime during the cDBS Baseline period will be used for analysis as subjects should be on stable cDBS stimulation from enrollment through the end of this period. In the aDBS Evaluation Phase, diary periods will be defined starting with the first completed diary record which is no more than 14 days prior to the visit. If diary is collected more than 14 days prior to the visit, the diary may be used for analysis if there is no change in aDBS stimulation settings. Complete diary periods are defined as any diary period where at least 21 hours (defined as 42 30-minute records) of diary were completed during the 24-hour period, and only complete periods will be used for analysis. If the diary is collected for more than 3 complete periods, up to 7 which are closest to the visit will be used for the analysis, while if a subject has zero complete periods, their diary will be considered missing. Diaries with at least 1 complete period available will be used for analysis. Data from non-contiguous complete periods may be used in the analysis.

The change in daily "On" time without troublesome dyskinesia from cDBS to aDBS will be computed for each subject. The overall standard deviation of the change will be computed for the FAS. The threshold will be set as 1 standard deviation. If the difference between aDBS and cDBS (aDBS-cDBS) for each subject is at least -1 standard deviation, the threshold will be met for the subject. Using this threshold, the proportion of subjects who exceed the threshold will be computed. The proportion with "On" time without troublesome dyskinesia during the respective aDBS Evaluation Period that exceed the threshold will be compared to the performance goal of 50% for each aDBS mode separately using a binomial exact test. The objective will be considered passed if the p-value is less than 0.0125. The lower bound of the 97.5% confidence interval will also be presented.

#### 7.9.1.5 Handling of missing data

All attempts will be made to minimize missing data. If 5% of the data or fewer are missing (as defined in section 7.9.1.4) for the analysis, no imputation will be used and the data will be analyzed as specified in the previous section. Otherwise, if missing data are observed, multiple imputation will be used for the analysis of the key primary objective.

Prior to the MI, the distribution of the continuous variables will be assessed for normality (using the Shapiro-Wilk test) to determine whether regression ($P \geq 0.05$) or predictive mean matching ($P < 0.05$) multiple imputation will be used for each variable. First, multiple imputation will be applied to impute missing values in continuous hours of "On" time without troublesome dyskinesia for each mode separately (aDBS single threshold, aDBS dual threshold). The model variables in the MI analysis may include, as appropriate:

- target site (STN or GPi)
- cDBS hours of "On" time without troublesome dyskinesia
- Age
- Gender
- Center
- PD duration

Variables will be ordered in the model based on the level of missingness (from least to most missing) and in the order specified above when the missingness is equal for 2 or more variables.

The fully conditional specification (FCS) method within SAS (version 9.4 or higher) with 5 repetitions and 100 burn-in iterations will be used for imputation. No limits will be defined for imputed values.

[redacted]

[redacted]

Following imputation, the "On" time without troublesome dyskinesia will be assessed against the "On" time threshold, and the estimate for the proportion of subjects meeting threshold will be determined for each imputation and combined using MI analysis methods. The lower bound of the 97.5% confidence limit will be determined using the MI Clopper-Pearson mean success approach[5]. In addition, the objective will be tested by comparing the lower bound of the 97.5% confidence limit to 50% for each imputation, and the overall p-value will be computed as the median of the p-values from the multiply imputed datasets.[6]

### 7.9.1.6 Supporting analyses

Supporting analyses for the primary objective in the Primary Cohort include verifying the lack of a period or carryover effect (randomized subjects only), a target site effect (STN/GPi), a lead type effect (legacy Models 3387 and 3389 compared to SenSight Models B33005 and B33015), or a center effect. Summarized results from the Directional Stimulation Cohort will be provided as supporting analyses to the Primary Cohort results. The threshold for meeting the primary objective in the Directional Stimulation Cohort will be the same threshold as calculated for the Primary Cohort. In addition, sensitivity analyses will be conducted using the CC and the AT Analysis Set.

An analogy of Prescott's test[7] will be computed to test for a period or carryover effect and p-values >0.1 will be deemed as supporting the assumption of no period or carryover effect in the proportion of subjects exceeding threshold for the primary outcome using data from both periods. If the assumption of no period or carryover effect is accepted, target site (STN/GPi), lead type, and center pooling assessments will be made using data from all subjects treated with aDBS in the aDBS Evaluation Phase. If a carryover effect exists, only data from the first period will be used for the analysis. [redacted]

[5] Mandan, Arpita (2018). *Two Applications of Summary Statistics: Integrating Information Across Genes and Confidence Intervals With Missing Data.* Master's thesis, Duke University. Retrieved from https://hdl.handle.net/10161/17533

[6] Ekhout, I., van de Wiel, M.A., and Heymans, M.W. (2017), Methods for significance testing of categorical covariates in logistic regression models after multiple imputation: power and applicability analysis. BMC Med. Res. Methodol.17: 129. doi: 10.1186/s12874-017-0404-7

[7] Nagelkerke, N.J.D., Hart, A.A.M. and Oosting, J. (1986), The Two Period Binary Response Cross - Over Trial. Biom. J., 28: 863-869. doi:10.1002/bimj.4710280715

- [redacted]
- [redacted]

To test for a treatment difference by target site (STN/GPi), lead type, and center within each mode, Fisher's exact test and Freeman and Halton's extension of Fisher's exact test, respectively, will be computed testing for differences between target sites, lead type, or centers in the proportion of subjects exceeding threshold for the primary outcome using data from both periods. [redacted]

- [redacted]
- [redacted]

P-values >0.1 will be deemed as supporting the assumption of no target site effect, lead type effect, or center effect, respectively, and the data will be considered poolable across the respective variable. If the p-value approaches statistical significance (defined as ≤0.1),[1] the proportion exceeding threshold and the lower bound of the 97.5% confidence interval will also be estimated at each target site, lead type, or center, as applicable (eg, if the target sites (STN/GPi) do not allow for pooling the analysis will be conducted by target site for that mode.)

If center(s) causing the significance are identified, variables relating to patient characteristics and other factors will be analyzed to try to identify why this center is showing a different treatment effect.

As part of the additional objectives, "On" time without troublesome dyskinesia will be summarized as absolute change from cDBS Baseline Phase to the Evaluation Phase.

## 7.9.2 Secondary Objective

### 7.9.2.1 Secondary Objective

To demonstrate decreased stimulation energy use during the aDBS Evaluation Phase as compared with cDBS.

### 7.9.2.2 Hypothesis

Two secondary hypotheses are being tested using independent gatekeeper procedures to control the alpha-level as described in section 7.9.1.1. The two secondary alternative hypotheses comparing aDBS (Evaluation Phase) to cDBS (Baseline Phase) separately, are:

- $H_{12}$: Mean Difference between aDBS dual threshold (Evaluation Phase) minus cDBS (Baseline Phase) for TEED < 0
- $H_{22}$: Mean Difference between aDBS single threshold (Evaluation Phase) minus cDBS (Baseline Phase) for TEED < 0

### 7.9.2.3 Endpoint definition

The secondary endpoint is to decrease energy use as measured by TEED during the aDBS Evaluation Phase as compared with TEED during the cDBS Baseline Phase.

TEED is defined as the total energy delivered by an electrical system through the DBS leads over an arbitrary period of time. TEED is determined by the stimulation parameters (e.g., pulse width, frequency, amplitude) and the measured impedance.

The calculation of TEED when using constant voltage is:

$$TEED = (Voltage^2 * Frequency * Pulse\ width) / Impedance$$

The calculation of TEED when using constant current is:

$$TEED = Current^2 * Frequency * Pulse\ width * Impedance$$

The Total TEED will be calculated as a sum of the energy use in the left and right leads. There will be one value per patient for each mode evaluated (eg, cDBS, aDBS Dual Threshold, aDBS Single Threshold).

The reduction in TEED will be calculated as:

$$TEED\ \Delta_{aDBS\ dual\ threshold\ mode} = TEED_{aDBS\ dual\ threshold\ mode} - TEED_{cDBS}$$

and

$$TEED\ \Delta_{aDBS\ single\ threshold\ mode} = TEED_{aDBS\ single\ threshold\ mode} - TEED_{cDBS}$$

#### 7.9.2.3.1 TEED calculations by lead type

*Calculation with legacy lead models 3387 or 3389*

The formulas in Section 7.9.2.3 can be used without modification.

*Calculation with SenSight™ system*

When SenSight is programmed in Directional Stim Mode or Ring Mode Stimulation, each contact is considered independent.
For each segment that is activated, TEED will be calculated individually using the formulas in Section 7.9.2.3, and then summed to obtain the total TEED for the lead.

#### 7.9.2.3.2 TEED calculation by visit

*Computing TEED in cDBS*

To compute TEED using the formula above in the cDBS mode, the current drain, pulse-width, frequency, and impedance will be determined from the JSON session data at the cDBS Baseline visit.

*Computing TEED in aDBS*

To compute TEED using the formula above in an aDBS mode, the average current drain from the JSON session data will be computed using the average stimulation currents from the timeline feature. This will be computed as the average of the currents from the timeline feature using the epochs that align with the analysis set (e.g., As-Treated analysis set uses the actual aDBS mode as reported from the device). Included days should be in the 14 days prior to the visit (ie, Visit 1 or Visit 2). The pulse-width, frequency, and impedance will be determined from the JSON session data at the respective aDBS Evaluation visit (ie, Visit 1 or Visit 2).

#### 7.9.2.3.3 TEED calculations by electrical stimulation mode

Leads programmed with monopolar settings will use the electrode unipolar impedance between the case and the active electrode. If 2 cathodes are used (dual monopolar settings), TEED will be calculated for each contact separately, and total TEED will be determined for the lead as the sum of the two TEED. Leads programmed to bipolar will use the impedance between the active electrodes. In case electrode impedance was not collected at cDBS Baseline or aDBS Evaluation Phase, impedance from aDBS Setup would be used.

### 7.9.2.4 Sample Size justification

The sample size was estimated using a one-sided test (alpha=0.0125) for one mean (PASS 2011 module Tests for One Mean) to compare the difference in aDBS TEED to cDBS TEED of 22 and a standard deviation of 35, a minimum of 36 subjects achieves at least 90% power.

Based on the registry data and the bench testing, a conservative estimate based on the first quartile of TEED data for initial devices in the registry, TEED is expected to decrease 35 and 22 units for aDBS single threshold mode and aDBS dual threshold mode, respectively. A SD of 36 results in a medium to large effect size of 0.629.

### 7.9.2.5 Analysis methods

The secondary analysis will use the FAS Analysis Set as described in section 7.1.3. Change in TEED will be used to evaluate the secondary objective. A paired t-test will be used to compare the difference in TEED for each aDBS mode (Evaluation Phase) to cDBS (Baseline Phase), separately. Normality of the outcome will be assessed with the Shapiro-Wilk test. If large violations to the normality assumption are observed ($p<0.05$), the primary analysis will use a Wilcoxon signed-rank test to assess statistical significance.

### 7.9.2.6 Handling of missing data

All attempts will be made to minimize missing data. If 5% of the data or fewer are missing for the analysis, no imputation will be used and the data will be analyzed as specified in the previous section. Otherwise, if missing data are observed, multiple imputation will be used for the analysis of the secondary objective.

Prior to the MI, the distribution of the continuous variables will be assessed for normality (using the Shapiro-Wilk test) to determine whether regression (P≥0.05) or predictive mean matching (P<0.05) multiple imputation will be used for each variable. First, multiple imputation will be applied to impute missing values in TEED for each mode separately (aDBS single threshold, aDBS dual threshold). The model variables to include and methods used for the MI analysis are described in Section 7.9.1.5, except that a variable for cDBS TEED will be included in the model instead of cDBS hours of "On" time without troublesome dyskinesia. The FCS method within SAS (version 9.4 or higher) with 5 repetitions and 100 burn-in iterations will be used for imputation. No limits will be defined for imputed values. [redacted]

Following imputation, the mean and standard error of the change in TEED will be determined for each imputation using Proc UNIVARIATE:

[redacted]

Proc MIANALYZE will be used to combine the summary statistics and to test the change in TEED with a t-statistic comparing the value to no change.

[redacted]

### 7.9.2.7 Supporting analyses

Supporting analyses for the secondary objective in the Primary Cohort include verifying the lack of a period or carryover effect (randomized subjects only), tests for poolability by target site (STN/GPi), tests for poolability by lead type (legacy Models 3387 and 3389 compared to SenSight Models B33005 and B33015), and tests for poolability by center. The test for a carryover effect will use a two-sample t-test,

all other supporting analysis tests will use linear regression models. In addition, results from the Directional Stimulation Cohort will be described as a supportive analysis. The test of no period or carryover effect will be computed first, and if the assumption of no period or carryover effect is accepted, data from both periods from all subjects treated with aDBS in the aDBS Evaluation Phase will be used for the models to assess poolability by target site, lead type, and center. The models will use the change between aDBS and cDBS and include an intercept and a covariate for the respective parameter.

If the P-value is >0.1, that will be deemed as supporting the assumption of no carryover effect and data from both periods from all subjects treated with aDBS in the aDBS Evaluation Phase will be used for the models to assess poolability by target site, lead type, and center. If a carryover effect exists (P-value ≤0.1), only data from the first period will be used for the analysis.

The tests for poolability of target site (STN/GPi), lead type, and center will be performed, by mode, and will include all subjects who enter the aDBS evaluation phase.

P-values >0.1 for the terms testing the effect of the respective parameters will be deemed as supporting the assumption of no target site effect, lead type effect, or center effect, respectively, and the data will be considered poolable across the respective variable. Normality of the outcome will be assessed with the Shapiro-Wilk test. If large violations to the normality assumption are observed ($p<0.05$), the

secondary endpoint may be transformed (e.g., log-transformation) or apply non-parametric method for the supporting analyses.

In addition, sensitivity analyses will be conducted using the CC and AT Analysis Sets without imputation.

### 7.9.3 Additional Objectives

To characterize each aDBS mode as compared to cDBS, the additional objectives described in sections 7.9.3.1 through 7.9.3.9 will be evaluated separately for each aDBS mode. The additional objectives in section 7.9.3.10 through 7.9.3.11 will characterize each aDBS mode, separately, without comparisons to cDBS. Additional objectives will be characterized for the CC unless otherwise specified, and the subgroup analyses described in section 4.4 will be performed on the CC analysis set and DSAT combined. Relevant references for additional objectives can be found in the study CIP.

Summary statistics may be presented for continuous measures (e.g., N, means, medians, standard deviations, minimums and maximums) and categorical measures (e.g., N, percentage, frequency distributions) with 95% confidence intervals as appropriate. If performed, missing data imputations will be conducted as specified within the respective sections. No adjustment will be made for multiple testing as no hypothesis testing will be conducted for the additional objectives.

To characterize aDBS as compared to cDBS for the objectives in sections 7.9.3.1 through 7.9.3.7, summary statistics will be provided in each aDBS mode and in cDBS along with the difference in aDBS and cDBS for continuous objectives. The difference will be calculated as follows for outcomes in which higher values are better:

$$\Delta_{\text{aDBS dual threshold mode}} = \theta_{\text{aDBS dual threshold mode}} - \theta_{\text{cDBS}}$$

and

$$\Delta_{\text{aDBS single threshold mode}} = \theta_{\text{aDBS single threshold mode}} - \theta_{\text{cDBS}}$$

And as follows for outcomes in which lower values are better:

$$\Delta_{\text{aDBS dual threshold mode}} = \theta_{\text{cDBS}} - \theta_{\text{aDBS dual threshold mode}}$$

and

$$\Delta_{\text{aDBS single threshold mode}} = \theta_{\text{cDBS}} - \theta_{\text{aDBS single threshold mode}}$$

Where Δ is the mean within-patient change in the objective between the respective aDBS mode and cDBS mode.

### 7.9.3.1 Voice Handicap Index (VHI)

The VHI has been designed to evaluate the quality-of-life specific to dysphonia. It was developed and validated with patients with a wide range of disorders, including neurological disorders and quantifies the patients' perceptions of the handicap they experience in everyday life due to speech disorders.

The 30-item self-administered questionnaire consists of a 10-item functional subscale, a 10-item emotional subscale, and a 10-item physical subscale. Subjects are asked to read each item and circle one of five response comprising an equal-appearing five-point scale. The scale has the words "never" and "always" anchoring each end and the words "almost never", "sometimes", and "almost always" appearing in between. An "always" response was scored 4 points, a "never" response was scored 0 points, and the remaining options were scored between 1 and 3 points. The score for each subscale is calculated by summing the scores within the subscale and the total score is calculated by summing the scores from the entire questionnaire.

The VHI will be summarized in both "best" and "worst" conditions. A subgroup analysis of this objective will also be performed for those with speech side effects as defined from the MDS-UPDRS question 3.1, score of 2 or greater (mild to severe) in the off medication, off stimulation state at the cDBS Baseline visit.

*Missing values*

If less than 50% of the items are missing from a subscale, the respective subscale score should be retained with the mean scale score of the items present used to impute a score for the missing items. The total score may be calculated by summing the imputed subscale scores. These imputations will be performed on the Complete Case analysis set.

### 7.9.3.2 Movement Disorders Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS)

The MDS-UPDRS is an MDS-sponsored revision of the UPDRS clinical rating scale for Parkinson's disease. It has been the most widely used scale to assess impairment and disability in PD. The MDS-UPDRS is a comprehensive 50 question assessment of both motor and non-motor PD symptoms. The assessment has four parts: Part I (non-motor experiences of daily living), Part II (motor experiences of daily living), Part III (motor examination) and Part IV (motor complications). Parts 1A, III and IV are assessed by the trained rater and Parts IB and II are a self-administered patient questionnaire that is completed with or without the aid of the caregiver.

Each question is scored from 0 to 4 reflecting the subject's usual level of function; Normal, Slight, Mild, Moderate and Severe. Each part (I-IV) can be reported as sub-scores, or all parts can be reported as an overall score.

These items will be summarized from the MDS-UPDRS:

- MDS-UPDRS III subscore (On stim/On med) of tremor, rigidity, and bradykinesia questions where a side can be determined: 3.3 b, c, d, e (rigidity), 3.4 a, b (finger taps/bradykinesia), 3.15 a, b

(postural tremor of hands), 3.16 a, b (kinetic tremor of hands), and 3.17 a, b, c, d (rest tremor amplitude)

- MDS-UPDRS III (On stim/On med)
- MDS-UPDRS II: "best" and "worst" conditions
- MDS-UPDRS III speech question (On stim/On med) q 3.1
- MDS-UPDRS IV

Additionally, for those with mixed programming (one hemisphere with aDBS and one with cDBS), MDS-UPDRS III will be summarized by side in the On stim/On med condition.

*Missing values*

Missing data will be imputed using the methods of Goetz et. Al.[8] The analysis will be performed on the CC analysis set when there is partial data meeting the requirements for imputation specified in Goetz et. Al. To provide valid part scores applicable across all Hoehn and Yahr (H&Y) stages when the same items are consistently missing, one missing item from Part I, one from Part II, three from Part III, but none from Part IV can be allowed. To provide valid part scores applicable across all H&Y stages when random item entries are missing, one missing item from Part I, two from Part II, seven from Part III, but none from Part IV can be allowed.

#### 7.9.3.3 European Quality of Life – 5 Dimensions (EQ-5D-5L)

The European Quality of Life – 5 Dimensions, version 5L (EQ-5D-5L), is a standardized measure of health status developed by the EuroQol Group and a widely used validated tool to determine health-related quality of life. The EQ-5D descriptive system consists of 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 5 levels of severity: no problems, slight problems, moderate problems, severe problems, extreme problems. The subject will be asked to indicate his/her health state by selecting the most appropriate statement in each of the 5 dimensions. This decision results in a 1-digit number expressing the level selected for that dimension. The digits for 5 dimensions can be combined into a 5-digit number describing the respondent's health state. It should be noted that the numerals 1-5 have no arithmetic properties and should not be used as a cardinal score. In a second step, the responses to the five EQ-5D dimensions (i.e. an EQ-5D health state or profile) will be converted into a single number called an index value using the United Kingdom value set. The index value reflects how good or bad the health state is according to the preferences of the general population of a country/region.

The EQ-5D dimensions will be summarized as categorical measures and the index value and visual analog scale (VAS) will be summarized as continuous measures. The 5-digit number describing the respondent's health state will not be summarized.

[8] Goetz, C.G., Luo, S., Wang, L., Tilley, B.C, LaPelle, N.R. and Stebbins, G.T. Handling Missing Values in the MDS-UPDRS. Mov Disord. 2015 Oct; 30(12): 1632-8, doi: 10.1002/mds.26153. Epub 2015 Feb 4.

*Missing values*

Missing data imputation will not be performed for the EQ-5D-5L.

### 7.9.3.4 Parkinson's Disease Sleep Scale 2 (PDSS-2)

The PDSS-2 assesses a wide spectrum of disease-specific sleep problems and is administered as a patient self-rating scale.

The PDSS-2 consists of 15 questions about various sleep and nocturnal disturbances which are to be rated by the patients using one of five categories, from 0 (never) to 4 (very frequent). PDSS-2 total score ranges from 0 (no disturbance) to 60 (maximum nocturnal disturbance).

*Missing values*

If less than 50% of the questions are missing, the PDSS-2 total score should be retained with the mean scale score of the items present used to impute a score for the missing items. These imputations will be performed on the Complete Case analysis set.

### 7.9.3.5 Wearable Data

Wearable data will be collected via the PKG Watch as a continuous measure of bradykinesia score (BKS) and dyskinesia score (DKS). A Fluctuation and Dyskinesia score (FDS) will also be derived from the wearable data. The analysis may include reliable data based on day or night on wrist count and all available data.

In addition to the main analysis of wearable data, a subgroup analysis will be performed for those with dyskinesias at least 25% of time. This will be defined from the MDS-UPDRS question 4.1, in the off medication, off stimulation state at the cDBS Baseline visit.

*Missing values*

Missing data imputation will not be performed with the wearable data.

### 7.9.3.6 Parkinson's Disease Questionnaire (PDQ-39)

The Parkinson's Disease Questionnaire is a self-completion Patient Reported Outcome (PRO) designed to address aspects of functioning and well-being for those affected by Parkinson's disease. There are 39 questions in the long form Parkinson's Disease Questionnaire, with 8 discrete domains:

- mobility (10 items, questions number 1 to 10)
- activities of daily living (6 items, questions number 11 to 16)
- emotional well-being (6 items, questions number 17 to 22)
- stigma (4 items, questions number 23 to 26)
- social support (3 items, questions number 27 to 29)
- cognitions (4 items, questions number 30 to 33)
- communication (3 items, questions number 34 to 36)
- bodily discomfort (3 items, questions number 37 to 39)

Patients are asked to think about their health and general well-being and to consider how often in the last month they have experienced certain events (e.g. difficulty walking 100 yards). Patients are asked to indicate the frequency of each event by selecting one of 5 options (Likert Scale): never/occasionally/sometimes/often/always or cannot do at all. *Never* is given a score of 0, and *always or cannot do* is given a score of 4, and the remaining options are scored between 1 and 3.

Each domain is calculated as a scale from 0 to 100, where 0 represents no problem at all and 100 represents the maximum level of problem. The formula for scoring each domain is:

$$\frac{\text{sum of scores of each question in domain}}{4 * \text{number of questions in domain}} \times 100$$

The PDQ-39 Summary Index (SI) and the 8 subscales will be summarized for this outcome[9]. In addition to the SI and its 8 subscales, the PDQ-39 speech questions 34 and 35 will be summarized individually as well.

The PDQ-39-SI is scored as:

$$\frac{\text{sum of all domain scores}}{8}$$

*Missing values*
If the response to a question is missing, no scale score is calculated for that individual for that domain. This precludes the calculation of the PDQ-39 summary index score.

### 7.9.3.7 Parkinson's Disease (PD) Home Diary

The Parkinson's Disease Home Diary is a standard home diary to assess functional status in patients with PD. In 30-minute intervals, patients record whether they were in the "On" condition, with or without dyskinesia, "Off" condition, or asleep. In the On condition with dyskinesia, patients record whether they were On with troublesome or non-troublesome dyskinesia.

These items will be summarized from the PD Home Diary:

- "On" time without troublesome dyskinesia
- "On" time without dyskinesia
- "On" time with non-troublesome dyskinesia
- "On" time with troublesome dyskinesia
- "Off" time

[9] Jenkinson, C., Fitzpatrick, R., Peto, V., Dunnett, S., Kelly, L. and Morley, D. The Parkinson's Disease Questionnaires User Manual, Fourth Edition. Health Services Research Unit, University of Oxford. 2018

- Asleep time

The analysis will use the FAS as described in Section 7.1.3. The PD Home Diary will be collected for at least 3 consecutive days (24 hour "periods") prior to each visit to evaluate the PD Home Diary additional objectives. Diary periods will be defined starting with the first completed diary record which is no more than 14 days prior to the visit. Complete diary periods are defined as any diary period where at least 21 hours (defined as 42 30-minute records) of diary were completed during the 24-hour period, and only complete periods will be used for analysis. If the diary is collected for more than 3 complete periods, up to 7 which are closest to the visit will be used for the analysis, while if a subject has zero complete periods, their diary will be considered missing. Diaries with at least 1 complete period available will be used for analysis. Included days cannot be more than 14 days prior to the visit. Data from non-contiguous complete periods may be used in the analysis.
The change in in daily averages from cDBS to aDBS will be computed for each item, for each subject. The overall standard deviation of the change will be computed for the FAS.

*Missing values*

If 5% of the data or fewer are missing (as defined in section 7.9.1.4) for the analysis, no imputation will be used and the data will be analyzed as specified in the previous section. Otherwise, if missing data are observed, multiple imputation will be used for the analysis of each item from the PD Home Diary.

Prior to the MI, the distribution of the continuous variables will be assessed for normality (using the Shapiro-Wilk test) to determine whether regression ($P \geqslant 0.05$) or predictive mean matching ($P<0.05$) multiple imputation will be used for each variable. First, multiple imputation will be applied to impute missing values in continuous hours of each item on the PD Home Diary for each mode separately (aDBS single threshold, aDBS dual threshold). Variables for target site (STN or GPi) and cDBS hours each applicable item will be considered for the model. The fully conditional specification (FCS) method within SAS (version 9.4 or higher) with 5 repetitions and 100 burn-in iterations will be used for imputation. No limits will be defined for imputed values. Following imputation, the change in in daily averages from cDBS to aDBS will be computed for each item, and the estimate and standard error of the change will be determined for each imputation. These will be combined into one result to estimate the change and standard error using MI analysis methods.

*Supporting analyses*

Sensitivity analyses will be conducted using the CC Analysis Set.

### 7.9.3.8 Patient Preference

Patient preference for aDBS vs. cDBS and Single vs. Dual Threshold aDBS modes (if both were evaluated) will be assessed by a Medtronic-developed questionnaire.

The questionnaire has three parts, the first assessing subject preference for aDBS as compared with the cDBS therapy they received before starting the study, to which subjects can reply that they strongly

prefer aDBS, somewhat prefer aDBS, no preference between aDBS and cDBS, somewhat prefer cDBS, or strongly prefer cDBS. The specific questions asked in part I are displayed in Figure 7-1.

The second part assesses subject blinding. Subjects report which aDBS mode they believe they received in the relevant period, why they believe that to be true, and how certain they are about their assessment. The specific questions are displayed in Figure 7-2.

The third part assesses subject preference for aDBS mode among subjects receiving therapy with both aDBS modes, to which subjects can reply that they prefer the first aDBS mode received, the second aDBS mode received, no preference, or do not know. The questions asked in part III are displayed in Figure 7-3.

**Figure 7-1. Evaluation of aDBS therapy as compared to DBS therapy received prior to study start**

**Section I: Evaluation of the DBS Therapy Received Over the Past 14 Days**

Instructions: We are interested in your evaluation of the deep brain stimulation (DBS) therapy you received over the past 14 days and would like you to compare it to the DBS therapy you were receiving before you joined the research study. Think about how effective the therapy was on your disease-related symptoms (examples: shaking, slowness, stiffness) and any side effects (examples: trouble swallowing, change in mood, tingling) caused by the DBS therapy you are receiving. For each question, please place a single check mark next to the response you feel best represents your experience and preferences.

**Overall**, do you prefer the DBS therapy you received during the past 14 days or the DBS therapy you were receiving prior to joining the research study

- o Strongly prefer the DBS therapy received during the last 14 days.
- o Somewhat prefer the DBS therapy received during the last 14 days.
- o No preference between the two.
- o Somewhat prefer the DBS therapy I was receiving prior to joining the research study.
- o Strongly prefer the DBS therapy I was receiving prior to joining the research study.

Based on your **improvement in movement symptoms** (*examples: slowness, stiffness, difficulty with speech, walking, balance, tremor*) alone, do you prefer the DBS therapy you received during the past 14 days or the DBS therapy you were receiving prior to joining the research study

- o Strongly prefer the DBS therapy received during the last 14 days.
- o Somewhat prefer the DBS therapy received during the last 14 days.
- o No preference between the two.
- o Somewhat prefer the DBS therapy I was receiving prior to joining the research study.
- o Strongly prefer the DBS therapy I was receiving prior to joining the research study.

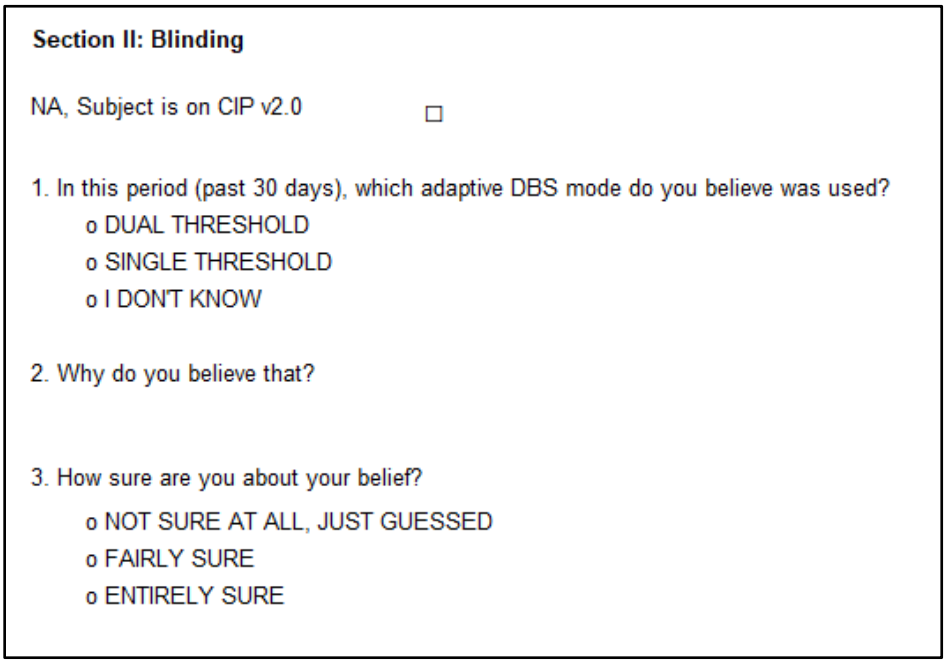
Based on **side effects** (*examples: trouble swallowing, change in mood, tingling*) alone, do you prefer the DBS therapy you received during the past 14 days or the DBS therapy you were receiving prior to joining the research study

- o Strongly prefer the DBS therapy received during the last 14 days.
- o Somewhat prefer the DBS therapy received during the last 14 days.
- o No preference between the two.
- o Somewhat prefer the DBS therapy I was receiving prior to joining the research study.
- o Strongly prefer the DBS therapy I was receiving prior to joining the research study.

Based on **symptom fluctuations** alone, do you prefer the DBS therapy you received during the past 14 days or the DBS therapy you were receiving prior to joining the research study

- o Strongly prefer the DBS therapy received during the last 14 days.
- o Somewhat prefer the DBS therapy received during the last 14 days.
- o No preference between the two.
- o Somewhat prefer the DBS therapy I was receiving prior to joining the research study.
- o Strongly prefer the DBS therapy I was receiving prior to joining the research study.

**Figure 7-2. Blinding**



**Section II: Blinding**

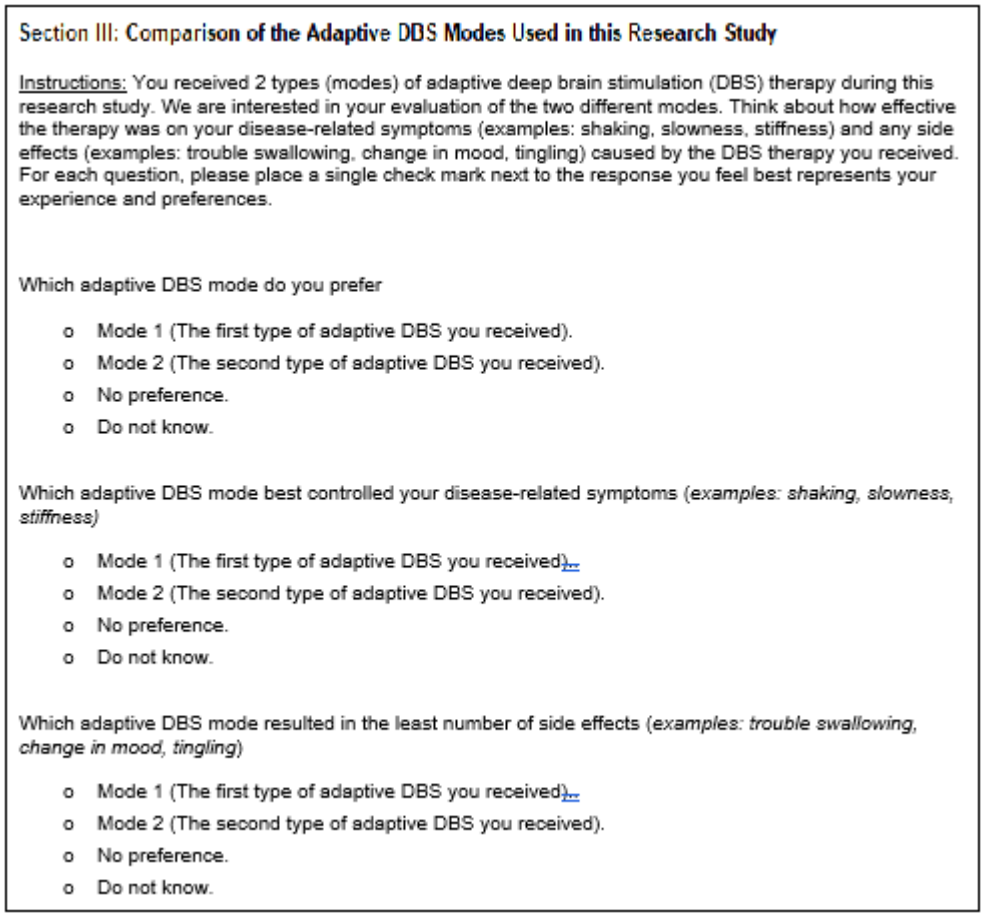
NA, Subject is on CIP v2.0 ☐

1. In this period (past 30 days), which adaptive DBS mode do you believe was used?
   - o DUAL THRESHOLD
   - o SINGLE THRESHOLD
   - o I DON'T KNOW

2. Why do you believe that?

3. How sure are you about your belief?
   - o NOT SURE AT ALL, JUST GUESSED
   - o FAIRLY SURE
   - o ENTIRELY SURE

**Figure 7-3. Comparison of aDBS modes**



Section III: Comparison of the Adaptive DBS Modes Used in this Research Study

Instructions: You received 2 types (modes) of adaptive deep brain stimulation (DBS) therapy during this research study. We are interested in your evaluation of the two different modes. Think about how effective the therapy was on your disease-related symptoms (examples: shaking, slowness, stiffness) and any side effects (examples: trouble swallowing, change in mood, tingling) caused by the DBS therapy you received. For each question, please place a single check mark next to the response you feel best represents your experience and preferences.

Which adaptive DBS mode do you prefer

- o Mode 1 (The first type of adaptive DBS you received).
- o Mode 2 (The second type of adaptive DBS you received).
- o No preference.
- o Do not know.

Which adaptive DBS mode best controlled your disease-related symptoms (*examples: shaking, slowness, stiffness)*

- o Mode 1 (The first type of adaptive DBS you received).
- o Mode 2 (The second type of adaptive DBS you received).
- o No preference.
- o Do not know.

Which adaptive DBS mode resulted in the least number of side effects (*examples: trouble swallowing, change in mood, tingling*)

- o Mode 1 (The first type of adaptive DBS you received).
- o Mode 2 (The second type of adaptive DBS you received).
- o No preference.
- o Do not know.

*Missing values*

Missing data imputation will not be performed with the patient preference objective.

### 7.9.3.9 Battery longevity

The Percept PC system has the ability to estimate the battery longevity based on programmed stimulation settings. This projected battery longevity will be determined from the JSON session data at each respective evaluation phase, and a subject will be considered a success if they have a projected battery longevity that is at least 25% greater within the respective aDBS visit as compared with the cDBS visit.

*Missing values*

Missing data imputation will not be performed for the battery longevity objective.

### 7.9.3.10 Patient Satisfaction

Patient satisfaction with aDBS during the long-term follow-up phase will be assessed by a Medtronic-developed questionnaire. Part I of the questionnaire is displayed in Figure 7-4, and Part II is displayed in Figure 7-5.

**Figure 7-4. Patient Satisfaction Questionnaire, Part I**

Instructions: We are interested in your satisfaction with the adaptive deep brain stimulation (DBS) therapy you received during the long term follow up phase of this study (approximately the past 9-10 months). Think about how effective the therapy was on your disease-related symptoms (examples: shaking, slowness, stiffness) and any side effects (examples: trouble swallowing, change in mood, tingling) caused by the DBS therapy you received. For each question, please place a single check mark next to the response you feel best represents how you would answer the question.

Would you recommend adaptive DBS to other DBS patients with Parkinson's Disease

- o Definitely yes.
- o Probably yes.
- o Neutral.
- o Probably not.
- o Definitely not.

Overall how satisfied or unsatisfied are you with the adaptive DBS you received during the long term follow up phase of this study

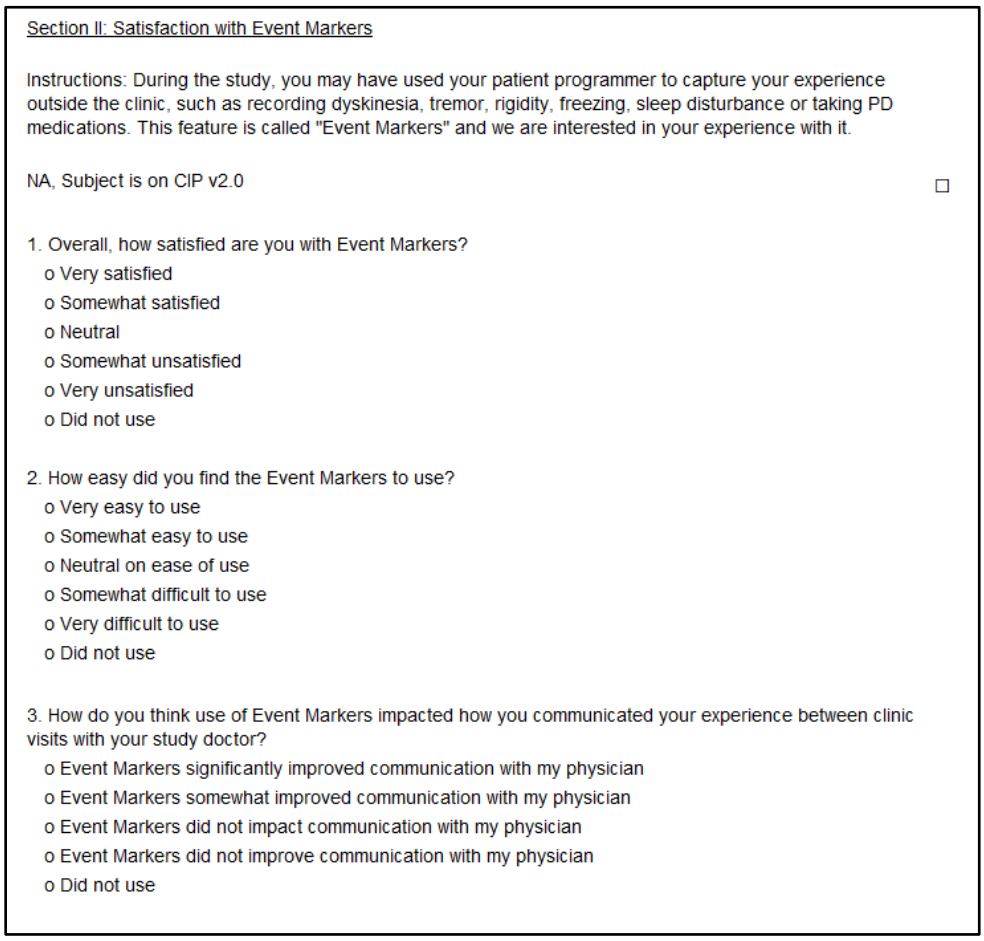
- o Very satisfied.
- o Somewhat satisfied.
- o Neutral
- o Somewhat unsatisfied.
- o Very unsatisfied.

What specific movement symptom was the most improved by adaptive DBS therapy

- o None
- o Speech
- o Chewing and swallowing
- o Handwriting
- o Tremor.
- o Walking and balance
- o Freezing
- o Stiffness / Rigidity
- o Slowness & hesitancy during movement / bradykinesia
- o Involuntary, random movements / dyskinesia
- o Posture
- o Cramps or spasms / dystonia
- o Other

If Other, Specify_____

**Figure 7-5. Patient Satisfaction Questionnaire, Part II**



Section II: Satisfaction with Event Markers

Instructions: During the study, you may have used your patient programmer to capture your experience outside the clinic, such as recording dyskinesia, tremor, rigidity, freezing, sleep disturbance or taking PD medications. This feature is called "Event Markers" and we are interested in your experience with it.

NA, Subject is on CIP v2.0 ☐

1. Overall, how satisfied are you with Event Markers?
   - o Very satisfied
   - o Somewhat satisfied
   - o Neutral
   - o Somewhat unsatisfied
   - o Very unsatisfied
   - o Did not use

2. How easy did you find the Event Markers to use?
   - o Very easy to use
   - o Somewhat easy to use
   - o Neutral on ease of use
   - o Somewhat difficult to use
   - o Very difficult to use
   - o Did not use

3. How do you think use of Event Markers impacted how you communicated your experience between clinic visits with your study doctor?
   - o Event Markers significantly improved communication with my physician
   - o Event Markers somewhat improved communication with my physician
   - o Event Markers did not impact communication with my physician
   - o Event Markers did not improve communication with my physician
   - o Did not use

*Missing values*

Missing data imputation will not be performed with the patient satisfaction objective.

### 7.9.3.11 Device data

The parameters summarized with the device data from the JSON file may include, but are not limited to:

- Active contacts
- Minimum and maximum voltage or current, per subject, for each aDBS Evaluation Phase
- Distribution of stimulation parameters
- Subjects tolerating aDBS for each aDBS mode
- Use of event markers in long-term follow-up
- aDBS on time
- aDBS Method parameters by visit
  - Decreasing ramp rate
  - Increasing ramp rate

    - Lower stimulation limit
    - Upper stimulation limit
- BrainSense data
    - LFP streaming data by therapy mode (cDBS, aDBS Single Threshold Mode, and aDBS Dual Threshold Mode)
    - LFP peak tracker
        - Peak Categories
            - Beta
            - Tremor
            - Gamma
            - Other
    - Peak amplitude
    - Peak frequency
    - Time spent below, between, and above thresholds for every group
    - Top LFP bipolar pair for each patient
- Circadian Rhythm tracker
    - Identify patients with circadian rhythms
    - Track the number of hours spent below daytime mean (7am-10pm) during nighttime (10pm-7am)

BrainSense data will be analyzed by the Medtronic Research and Technology group, or designee. Key BrainSense analyses will be presented in the final clinical study report.

A subgroup analysis will be performed to characterize the subject population and satisfaction with the use of event markers among subjects who use event markers during long-term follow-up.

*Missing values*

Missing data imputation will not be performed for the device data objective.

## 7.10 Safety Evaluation

The safety assessment is to characterize:

- Stimulation-related adverse events during the aDBS evaluation and the cDBS baseline phases.
- Serious adverse events, adverse events, and device deficiencies throughout the study.

Stimulation-related adverse events will be summarized during the cDBS Baseline and aDBS Evaluation Phases for all subjects, with tabulations individualized by aDBS mode in the Primary Cohort, by target site in the Primary Cohort, by lead type in the Primary Cohort, and the Directional Stimulation Cohort (DSAT analysis set). Summarized results from the Directional Stimulation Cohort will be provided as supporting analyses to the Primary Cohort results. In the Primary Cohort, if statistical comparisons of poolability across mode, lead type, and target site are accepted, overall tabulations will be provided as appropriate. To test for a difference in safety by target site (STN/GPi) and lead type, Fisher's exact test will be computed testing for differences between target sites and between lead type in the proportion

of subjects with a stimulation-related adverse event during the aDBS Evaluation Phase using data from both periods. [redacted]

A repeated measures logistic regression will be used to test for a difference between aDBS modes in the proportion of subjects with stimulation-related adverse events in the randomized subjects who enter the aDBS evaluation phase in the Primary Cohort. [redacted]

P-values >0.1 will be deemed as supporting the assumption of no target site, lead type, or mode effect, respectively, and the data will be considered poolable across the respective variable. If the p-value approaches statistical significance (defined as ≤0.1),[1] safety summaries will not be provided as a combined summary for that variable.

Serious adverse events, adverse events, and device deficiencies that occur from the time of enrollment through study exit/discontinuation will be summarized for all enrolled subjects. Analyses will be conducted on the AT analysis set for subjects in the Primary Cohort and on the DSAT for subjects in the Directional Stimulation Cohort. The denominator will include the appropriate number of subjects for each respective phase, period, and cohort. Events will be summarized by number of events, number of subjects who experienced the event, and percentage of subjects who experienced the event.

## 7.11 Changes to Planned Analysis

Any change to the data analysis methods described in the SAP will require an amendment only if it changes a principal feature of the SAP. Any other change to the data analysis methods described in the SAP, and the justification for making the change, will be described in the clinical study report.

# 8. Validation Requirements

Statistical programming code that affects the result of the main analysis (e.g., not including sensitivity or supporting analyses) for the primary objective shall be validated using Level I validation. Programming code for randomization and programming code that affects the result of the main analysis for the secondary objective(s) shall be validated using at least Level II validation. In addition, those main statistical analyses that are planned for publication and have not been previously validated should be validated using at least Level II validation. The CIP deviation summary shall be validated using at least Level III validation and the high-level adverse event summary shall be validated using at least Level II validation. Additional measures where a p-value or confidence interval has been generated may need to be validated using at least Level II validation.

# 9. References

[1] Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res. 1993; 2(2):121-45.

[2] Buhi, E, Goodson, P, Neilands T. Out of Sight, Not Out of Mind: Strategies for Handling Missing Data. Am J Health Behav. 2008 Jan-Feb;32(1):83-92, doi: 10.5555/ajhb.2008.32.1.83.

[3] Rubin, D. B. (1976). "Inference and Missing Data." Biometrika 63:581–592.

[4] Rubin, D. B. (1987). Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons.

[5] Mandan, Arpita (2018). *Two Applications of Summary Statistics: Integrating Information Across Genes and Confidence Intervals With Missing Data.* Master's thesis, Duke University. Retrieved from https://hdl.handle.net/10161/17533

[6] Ekhout, I., van de Wiel, M.A., and Heymans, M.W. (2017), Methods for significance testing of categorical covariates in logistic regression models after multiple imputation: power and applicability analysis. BMC Med. Res. Methodol.17: 129. doi: 10.1186/s12874-017-0404-7

[7] Nagelkerke, N.J.D., Hart, A.A.M. and Oosting, J. (1986), The Two Period Binary Response Cross - Over Trial. Biom. J., 28: 863-869. doi:10.1002/bimj.4710280715

[8] Goetz, C.G., Luo, S., Wang, L., Tilley, B.C, LaPelle, N.R. and Stebbins, G.T. Handling Missing Values in the MDS-UPDRS. Mov Disord. 2015 Oct; 30(12): 1632-8, doi: 10.1002/mds.26153. Epub 2015 Feb 4.

[9] Jenkinson, C., Fitzpatrick, R., Peto, V., Dunnett, S., Kelly, L. and Morley, D. The Parkinson's Disease Questionnaires User Manual, Fourth Edition. Health Services Research Unit, University of Oxford. 2018