Study Title: Clinical Study of Neuspera's Implantable Sacral Nerve Stimulation (SNS) System in Patients with Symptoms of Urinary Urgency Incontinence (UUI)

Document: Clinical Investigation Plan / Study Protocol

Protocol No. NSM-004
Version 21.0
04 Apr 2025

NCT number: NCT04232696

Neuspera

Clinical Study of Neuspera's Implantable Sacral Nerve Stimulation (SNS) System in Patients with Symptoms of Urinary Urgency Incontinence (UUI)

Protocol No. NSM-004, Version 20.0

Statistical Analysis Plan

Version 17.0

16 MAR 2024

Version History

Version	Version Date	Author/Title	Summary of Key Changes
1.0	21May2019	Chris Mullin/Statistician NAMSA	Initial Release
2.0	22Jul2019	Chris Mullin/Statistician NAMSA	Refined analysis populations and added performance goal discussion
3.0	30Oct2019	Chris Mullin/Statistician NAMSA	Updates made to be in alignment with protocol version 5.0 list of secondary endpoints.
4.0	05Dec2019	Chris Mullin/Statistician NAMSA	Updated NSM-004 protocol reference in document to most current version of protocol (V6 dated 05 December 2019)
5.0	25Feb2020	Chris Mullin/Statistician NAMSA	Updated NSM-004 protocol reference in document to most current version of protocol (V7 dated 28 January 2020)
6.0	8 May 2020	Chris Mullin/Statistician NAMSA	Updated NSM-004 protocol reference in the document to most current version of protocol (V9 dated 20 May 2020)
7.0	30 July 2020	Chris Mullin/Statistician NAMSA	Updated NSM-004 protocol reference in the document to most current version of protocol (V10 dated 17 July 2020)
8.0	06 Oct 2020	Kaisa Kivilaid/Statistician Covance	Updated NSM-004 protocol reference in the document to the most current version of protocol (V11.0 dated 06 October 2020). Changed number of subjects in Phase I to may be up to a total of 55, as indicated in protocol V11.0. Added section defining enrollment.

Version	Version Date	Author/Title	Summary of Key Changes
9.0	15 January 2021	Kaisa Kivilaid/Statistician Covance	Added analysis plan for hierarchical testing of secondary endpoints intended for labeling claims. Added null and alternative hypotheses for secondary endpoints intended for labeling claims. Updated definitions of analysis populations. Updated sample size. Updated design to remove references to trial period. Follow-up period begins after implant.
10.0	16 March 2021	Kaisa Kivilaid/Statistician Covance	Updated definitions of analysis populations. Elaborated on treatment of missing data for the primary and secondary efficacy endpoints intended for labeling claims.
11.0	16 July 2021	Jeffrey Verdoliva Boatman/Statistician LabCorp	Added references for clinical justification of efficacy endpoint. Added Monitoring of Phase II Safety Data for DSMB review.

Version	Version Date	Author/Title	Summary of Key Changes
12.0	17 November 2021	Jeffrey Verdoliva Boatman/Statistician LabCorp	Clarified that physician and subject satisfaction secondary endpoints are evaluated at 6 months. Updated surgical revision rates and added references. Added Phase II stopping rules and rational for stopping rules. Added formal mathematical statements for probabilities that Phase II event rates exceed comparator event rates.
13.0	08 March 2022	Jeffrey Verdoliva Boatman/Statistician Labcorp	Updated Phase II stopping rules for continual monitoring of device fracture, dislodgement/migration, and surgical revisions. Updated actions taken if a stopping rule is met. Updated secondary endpoints to match exactly the protocol synopsis and Statistical Consideration section of the protocol.
14.0	26 April 2022	Anna Nordell/Statistician NAMSA	Updated Phase II study design to group sequential, allowing for an analysis of the primary efficacy endpoint when 50% of subjects have completed 6-months of follow-up. Added language regarding a sensitivity analysis of the primary efficacy endpoint that will be performed accounting for subjects that used OAB medication during study follow-up

Version	Version Date	Author/Title	Summary of Key Changes
15.0	7 April 2023	Mark Jaros, PhD Statistician Summit Analytical	• Updated SAP based on changes made to protocol V18 (05APR2023). • Added new Intent-to-Treat (ITT) analysis set and updated original ITT analysis set to Modified ITT analysis set. • Added new secondary efficacy endpoint (change in urgent voids per day). • Clarified that secondary efficacy endpoints may be used for labeling claims instead of intending to be used for labeling claims. • New sequential testing order for secondary endpoints that may be used for labeling claims. • Updated interim analysis section.

Version	Version Date	Author/Title	Summary of Key Changes
16.0	12 June 2023	Mark Jaros, PhD Statistician Summit Analytical	• Updated SAP based on changes made to protocol V19 (12June2023) • Noted that no more than 5 sites will be included from Europe and Canada, and more than 50% of subjects enrolled will be from the United States. • Lowered the number of attempted implants from 145 to 120. • Lower the power for a one-sided 0.025 exact binomial test from 90 to 85%. • Increased number of clinical sites from 25 to 35. • Updated the "worst case" scenario results that would still meet the statistical threshold for this performance goal, changed to 72 out of 120 (60%). • Updated the assumed population responder rate from 64% to 63.5%. • Included site and region (US vs. OUS) as subgroups. • Added a ICIQ-OABqol responder rate secondary endpoint to be potentially used for labeling claims. • Updated information on PMA submission to match protocol. • Updated poolability analyses. • Updated details on interim analysis. • Added Coyne, et al reference and renumbered Bibliography

Version	Version Date	Author/Title	Summary of Key Changes
17.0	16 Mar 2024	Kaisa Kivilaid/Statistician Fortrea	• Update language regarding failures for primary efficacy endpoint to explicitly count explants/withdrawals and subjects taking OAB medications as failures • Note that subgroup analysis is exploratory and not to be used for labeling purposes • Refine secondary endpoints per FDA feedback; clarify analysis of ICIQ-OABqol is to be calculated according to published criteria

Document Approval

Approval	
Approved By	Kaisa Kivilaid
Signature	
Date	
Role	Biostatistician, Fortrea

Approval	
Approved By	Courtney Lane, PhD
Signature	
Date	
Role	VP, Clinical, Quality, and Regulatory Affairs

Table of Contents

1 Introduction

This statistical analysis plan describes the planned statistical methods to be used during the reporting and analysis of data collected under the Clinical Investigation Protocol, "Clinical Study of Neuspera's Implantable Sacral Nerve Stimulation (SNS) System in Patients with Symptoms of Urinary Urgency Incontinence (UUI)". This SAP should be read in conjunction with the study Clinical Investigation Plan and associated case report forms. This version of the analysis plan has been developed with respect to the Study Protocol Version 20.0, 16 Mar 2024. Any revisions to the protocol or CRFs that impact the planned analyses may require updates to the SAP.

2 Study Objectives

The overall objective of the study is to assess the safety and efficacy of the Neuspera Implantable SNS System for treatment of Urinary Urgency (UUI) based on Phase II results, following the determination of daily stimulation settings as informed by Phase I. Accordingly, hypotheses and formal endpoints are primarily based on results of Phase II. Unless otherwise specified, all analyses described in this section refer to the analysis of data gathered only under Phase II. Analyses for Phase I will consist of descriptive statistics since stand-alone definitive conclusions regarding safety or efficacy from this phase cannot be concluded due to the small sample size. The purpose of Phase I is simply to inform stimulation duration for Phase II.

3 Study Design

The study is a prospective, multi-center, single-arm, seamless phased pivotal study. The study population consists of subjects diagnosed with UUI who have failed or could not tolerate more conservative treatment.

The study will be conducted in two phases: Phase I may enroll up to 55 subjects in order to assess the utilization of the system during the SNS trial period and to help inform the length of hours of daily stimulation to be used in the second phase; Phase II may enroll up to 255 subjects (a separate group from Phase I subjects) to assess the safety and efficacy of the Neuspera SNS system at 6 months for the primary efficacy endpoint, and again at 12 months for the secondary safety and efficacy endpoints. The study is an operationally seamless design (in contrast to an inferentially seamless design). Accordingly, the formal statistical conclusions are to be based on solely Phase II. As there is no overlap of subjects, and the design of Phase II is fixed (i.e. there is no sample size re-estimation or other adaptive design elements), there are no bias or type I error concerns.

Phase I analysis may occur at several points during the trial as the sponsor wishes to assess compliance and the performance of the system as this will be the first chronic use of the system (See section 5.10 Interim Analysis). An interim analysis of Phase I data will occur at the discretion of the sponsor. The data from Phase I will not be the basis for formal conclusions; the intent is to gain initial data to inform the stimulation duration for Phase II.

A primary efficacy endpoint analysis of Phase II will occur when all subjects have completed 6 months post-implant (or are exited from the study). An interim analysis will be performed when the 60th Phase II subject has completed the 3-month post implant visit.

Sections 5.7 and 5.10 provide additional details on the timing of the analyses and planned interim analyses, respectively. All subjects are anticipated to be enrolled and implanted prior to the Phase II interim analysis. The Phase II interim analysis will be conducted by an independent statistician.

Regardless of the outcome of the interim analysis, the study will enroll to its planned sample size and all subjects will complete their scheduled follow-up (unless they are exited from the study). If the interim analysis is not successful, but the primary endpoint is achieved at the end of six months of follow-up, the sponsor intends to file the PMA submission.

It is expected that no more than 5 sites to be included from Europe and Canada. More than 50% of the subjects enrolled will be from the United States.

4 Sample Size Determination

Sample size for Phase I is not based on power requirements for a statistical hypothesis test, but for a desire to obtain a reasonable initial sample size to inform Phase II in a timely fashion. It is anticipated that approximately 33% of implanted subjects will not respond to SNS therapy during the trialing period or drop out for other reasons. It is also expected that some subjects may not demonstrate appropriate motor response during intra-operative testing. Therefore, to account for subjects in whom implant is attempted but the system is not implanted, up to 55 subjects may need to be enrolled to implant up to 40 subjects for 20-27 subjects to proceed post-trialing period in Phase I.

Sample size for Phase II is based on power requirements for the primary efficacy endpoint. A total of 120 attempted implants will provide 85% power for a one-sided 0.025 exact binomial test against the null hypothesis assuming that 1) the implant success rate is 95%, 2) subjects in whom implants are attempted but are not successful are counted as failures, and 3) the overall responder rate, including failed implant attempts, is 63.5%. Calculations are based on a one-sample, one-sided exact test of binomial proportions with a normal approximation and were performed with PASS 2019. To account for

potential attrition prior to implant of up to approximately 43% including screen failures, up to 255 subjects will be enrolled/consented in Phase II. Revisions will not count towards attempted implants. Phase I of the study will be conducted at up to 14 clinical study sites in the US and Europe. Phase II of the study will be conducted in up to 35 clinical sites in the US and Europe. Sample size at any one investigational site will be restricted to no more than 15 subjects enrolled for Phase I and 25 subjects implanted for Phase II. Consideration will be given to lifting the 25 subject implanted cap in Phase II if the investigator makes a request of the sponsor.

5 Statistical Analyses

5.1 General Considerations

Except where otherwise specified, the following general principles apply to the planned statistical analyses. All statistical analysis will be conducted using SAS version 9.4 or later (SAS Institute Inc., Cary, NC) or other widely accepted statistical or graphical software as required.

Unless otherwise noted, analyses will be performed separately for Phase I and II, with the focus of analyses on Phase II. Statistical tests are only planned based on Phase II. Any p-values based on Phase I will be considered exploratory.

5.1.1 Descriptive Statistics

Continuous data will be summarized with mean, standard deviation, median, minimum, maximum, and number of evaluable observations. Categorical variables will be summarized with frequency counts and percentages. Confidence intervals may be presented, where appropriate, using the t-distribution for continuous data and Clopper-Pearson Exact method for categorical variables.

5.1.2 Study Day 0/Baseline Assessments

Efficacy assessments that are measured over time will use pre-implant data as baseline.

For safety assessments, study day 0 is based on the date of implant (or date of attempted implant for those subjects attempted but not implanted, or date of replacement for subjects undergoing a replacement prior to the assessment of primary endpoint). Time of the event is based on the onset date for the event and calculated from date of implant (or date of attempted implant, or date of replacement, as appropriate). Specific time points for evaluation of safety events (e.g. 6 months) will be defined based on the corresponding visit date for subjects with a visit, or the close of the corresponding visit window for subjects without an event.

5.1.3 Visit Windows

Unless otherwise specified, visit-based assessments will be analyzed for each analysis time point according to the nominal visit entered in the Case Report Form (CRF).

5.1.4 Statistical Significance

Unless otherwise specified, hypothesis testing will be performed at the two-sided 0.05 significance level. P-values will be rounded to three decimal places. If a p-value is less than 0.0001 it will be reported as "<0.0001".

5.2 Analysis Populations

The following analysis populations are defined for analysis of subjects:

1. ***Intent-to-Treat (ITT) Analysis Set****: The ITT analysis set is defined as all Phase II subjects who have signed informed consent. Demographic information and study exit details will be reported separately for those subjects included in the ITT population where no initial implant was attempted.*
2. ***Modified Intent-to-Treat (mITT) Analysis Set****: The mITT analysis set is defined as all Phase II subjects for whom an initial implant is attempted. The mITT analysis set is the primary analysis set for assessment of primary endpoints.*
3. ***Implanted Analysis Set****: The Implanted analysis set is defined as all Phase II subjects who are implanted with the Neuspera SNS system.*
4. ***Responder Analysis Set****: The Responder analysis set is defined as subjects who are implanted with the system and are defined as responders at the Day 28 or Day 42 clinic visit post-implant.*
5. ***Per Protocol (PP) Analysis Set:*** *The Per Protocol analysis set is defined as implanted subjects who have no protocol deviations that may significantly affect the primary endpoints. The determination of subjects excluded from Per Protocol analysis set will be made prior to analysis.*

5.3 Subject Disposition

The number of subjects in each analysis population will be presented along with reason for any exclusions. Subject accountability will be summarized by visit. The number of subjects who are enrolled, attempted but not implanted, implanted, responders on Day 28 and Day 42, and number completing each clinical follow-up visit will be summarized. In addition, the number of subjects who complete the study or exit early will be summarized by reason.

5.4 Enrollment

During Phase I, subjects will be considered enrolled in the study after signing the informed consent form and after an implant is attempted. The principal investigator will ensure that the subject has met all pre-operative eligibility criteria prior to enrollment in the study. The intraoperative eligibility criterion of an appropriate motor response is confirmed during the implant procedure. Subjects in whom an implant is attempted but the appropriate motor response is not observed, will be considered enrolled in the study but will not go on to receive the neurostimulator implant.

During Phase II, subjects will be considered enrolled at the time the subject signs and dates the study specific informed consent form. Subjects who do not meet all eligibility criteria are considered enrolled in the study but will not receive an implant. Subjects in whom an implant is attempted but the appropriate motor response is not observed, will be considered enrolled in the study but will not go on to receive the neurostimulator implant.

5.5 Demographics and Baseline Characteristics

Descriptive statistics will be presented for all clinically relevant baseline demographic, medical history, and clinical characteristic variables.

5.6 Analysis of Study Endpoints

5.6.1 Primary Safety Endpoint

The primary safety endpoint is defined as the incidence of device-related SAEs through the 6-month post-implant follow-up visit. The analysis of the endpoint is the proportion of subjects experiencing a device-related SAE through the 6-month post-implant follow-up.

5.6.1.1 Primary Safety Analysis

There is no planned statistical hypothesis test for the primary safety endpoint. The proportion of subjects experiencing an event, as well as the numerator and denominator, and the associated two-sided exact 95% binomial confidence interval will be provided. Values will be expressed as percentages.

The primary analysis population for the primary safety endpoint is the mITT Analysis set as defined in Section 5.2. Subjects who exit the study prior to reaching the 6-month post-implant follow-up visit without experiencing a device-related SAE will be assumed free of such an event. Any adverse events reported for enrolled subjects who are excluded from mITT Analysis set will be reported separately for the duration of participation in the study.

5.6.2 Primary Efficacy Endpoint

The primary efficacy endpoint is defined as the percentage of subjects who experience an improvement in UUI symptoms of 50% or more (therapy responders). A therapy responder is defined as experiencing ≥ 50% reduction in the number of UUI episodes at 6 months post-implant, relative to the number of UUI episodes at baseline. For the primary efficacy endpoint, subjects withdrawing for reasons related to a lack of efficacy or due to an adverse event (including death) prior to the 6-month visit will be counted as failures. Those subjects who take OAB medications (after the first two weeks following implant) and prior to the primary efficacy endpoint shall also be considered as device failures.

Calculation of the primary efficacy endpoint will be based on the mITT Analysis Set as defined in Section 5.2.

5.6.2.1 Primary Efficacy Analysis

The statistical hypothesis test for the primary efficacy endpoint will be based on comparing the percentage of subjects with at least a 50% improvement in UUI symptoms (therapy responders) to a null value of 50%, the same performance goal as in the recent study by McCrery et al.[i] Mathematically, the null and alternative hypothesis are stated as:

$H_o: p \leq 0.50$
$H_a: p > 0.50$

The null hypothesis will be rejected at the 0.025 alpha level based on a one-sided, one sample exact tests of binomial proportions for responders. The proportion of 0.50 for responders corresponds to a value of 50%. Study success is defined as successful rejection of the null hypothesis.

The 50% reduction in UUI symptoms treats subjects as their own control and is based on a clinically meaningful improvement at the subject level in the primary efficacy endpoint[i,ii,iii]. Successful rejection of the null hypothesis will indicate a clinically meaningful result in the majority of subjects by a statistically significant amount. Success for the statistical test will require that the lower confidence bound for the percentage of subjects with at least 50% improvement in UUI symptoms (therapy responders) will exceed 50%. This means practically, that the percentage of subjects with at least a 50% improvement must be substantially greater than 50% in order for the lower confidence bound to exceed 50%. For example, with an evaluable sample size of 145 subjects, the "worst case" results that would still meet the statistical threshold for this performance goal would be 72 out of 120 (60%). Additionally, the study is powered around an assumed population responder rate of 63.5% or higher.

The primary analysis population for the primary efficacy endpoint is the mITT Analysis Set as defined in Section 5.2. Section 5.6.4 describes missing data handling for the primary efficacy endpoint in the mITT analysis set.

The primary efficacy endpoint will also be evaluated in the ITT, Implanted analysis set, Responder analysis set, and the Per Protocol analysis set (without imputation for missing data) to assess the robustness of results. If there are no differences between analyses sets (e.g., Implanted and Responder analysis sets include all the same subjects), separate analyses will not be presented. Analyses of the primary efficacy endpoint for all analysis populations will be conducted after 6-month post-implant follow-up is completed for all subjects.

The primary efficacy endpoint will be presented by study visit with descriptive statistics and nominal 95% confidence intervals.

5.6.2.2 Additional Analyses

Subgroup analysis of the primary efficacy endpoint will be performed for the following subgroups: sex, age (Age<65, Age≥65), race (white vs. non-white), baseline UUI episodes, site, and region (US vs. OUS). These analyses are intended to demonstrate consistency of results across subgroups. At each level of the subgroup of interest, the within group estimates will be summarized with frequency counts and percentages. The subgroup analysis of the primary efficacy endpoint are regarded as exploratory and will not be used for labeling purposes and will be summarized with descriptive statistics.

For each subgroup, a logistic regression model will be fit that includes a fixed effect for subgroup membership. If the p-value for the subgroup term is less than 0.15, additional exploratory analysis may be performed to understand any variations in outcomes by subgroup. These analyses will be performed using the mITT Analysis Set as defined in Section 5.2, but will include available data only (with no imputation planned for missing outcome values).

Further, as a sensitivity analysis of the primary efficacy endpoint, subjects that start an OAB medication after sacral nerve stimulation and prior to the 6-month follow-up visit will be included in the analysis according to their actual outcomes, and the efficacy results will be summarized. Additional sensitivity analyses of the primary efficacy endpoint may be performed if deemed of interest. Such analyses will be considered exploratory.

5.6.3 Secondary Endpoints

The secondary endpoints completed at 6 and 12 months if not indicated differently are below:

- Change from baseline in quality of life as measured and assessed by the total ICIQ-OABqol score.
- Change from baseline in mean number of UUI episodes.
- The percentage of subjects who experience an improvement in UUI symptoms (therapy responders) of at least 50% or more at 12 months visit post implant.
- The percentage of subjects who experience an improvement in ICIQ-OABqol of at least 10 points.
- Change in urgent voids per day calculated across all diary episodes with at least mild urgency.
- Change in average number of daily voids from baseline in subjects with at least 8 voids at baseline.
- Change in quality of life measured from baseline as measured and assessed by the ICIQ-OABqol subscale scores.
- Comprehensive summary of all adverse events (AEs) for the duration of study participation.
- Device parameters including but not limited to voltage, pulse width, frequency, and stimulating electrode.
- Physician and subject satisfaction as assessed with the User Experience Questionnaire at 6-month visit post implant.
- Change in Male/Female Lower Urinary Tract Symptoms questionnaire.
- The percentage of subjects with device-related serious adverse events reported through 12-month visit post-implant.
- Change in total urinary output as measured by 72-hour bladder diary.
- Change in fecal incontinence as measured by the Wexner Scale compared to baseline. Calculated in subjects with fecal incontinence.
- Patient Global Impression of Improvement (PGI-I) measured after implant during follow-up.

5.6.3.1 Analysis of Secondary Endpoints That May be Used for Labeling Claims

The primary analysis population for the secondary efficacy endpoints that may be used for labeling claims is the mITT analysis set as defined in Section 5.2. Section 5.6.4 describes missing data handling for these secondary efficacy endpoints in the mITT analysis set. These endpoints may also be assessed in other analysis sets without imputation for missing values.

The type I error for the primary efficacy endpoint and for a subset of secondary endpoints will be controlled via hierarchical testing, in the order listed below.

The following secondary endpoints will be included in a hierarchical analysis of study endpoints and may be intended for labeling claims (all at 6-month visit post-implant follow-up):

- Change from baseline in mean number of UUI episodes
- Change from baseline in ICIQ-OABqol score
- The percentage of subjects who experience an improvement in ICIQ-OABqol of at least 10 points
- Change in urgent episodes per day from baseline. Calculated across all diary episodes with at least mild urgency
- Change in average number of daily voids from in subjects with at least 8 voids at baseline
- Change in fecal incontinence as measured by the Wexner Scale compared to baseline. Calculated in subjects with fecal incontinence.

For the ICIQ-OABqol, analysis will be conducted according to published scoring criteria (Coyne et al., 2002)[iv].

The statistical hypothesis tests for continuous endpoints will be based on comparing the mean change scores at 6 months or 12 months to 0. Mathematically, the null and alternative hypotheses are stated as:

$H_o: \mu \leq 0$
$H_a: \mu > 0$

The statistical hypothesis test for the categorical endpoint based on UUI episodes will be based on comparing the percentage of subjects with a 50% improvement in UUI symptoms (therapy responders) at 12 months to a null value of 50%. Mathematically, the null and alternative hypothesis are stated as:

$H_o: p \leq 0.50$
$H_a: p > 0.50$

The statistical hypothesis test for the categorical endpoint based on ICIQ-OABqol will be based on comparing the percentage of subjects with at 10 point improvement (the minimally clinically important

difference, MCID; Coyne et al., 2006)[v] at 6 or 12 months to a null value of 50%. Mathematically, the null and alternative hypothesis are stated as:

Ho: $p \leq 0.50$
Ha: $p > 0.50$

The primary endpoint will be analyzed with a one-sided test at the $\alpha = 0.025$ level. To maintain the overall type I error rate, testing of these secondary hypotheses will only proceed if the analysis for the primary efficacy endpoint is statistically significant (one-sided $p < 0.025$. If the primary efficacy analysis is not statistically significant (one-sided $p \geq 0.025$), these secondary endpoints will not be formally tested, and analysis of these endpoints will be considered exploratory.
Provided that the primary endpoint analysis is statistically significant, testing will proceed to the secondary endpoints included in the hierarchical analysis. Testing will proceed in the order specified above with each endpoint tested sequentially. If the analysis of the secondary endpoint under consideration is statistically significant (one-sided $p < 0.025$, the result for that endpoint is declared statistically significant, and testing proceeds to the secondary endpoint next in the order. As soon as an analysis is not statistically significant (one-sided $p \geq 0.025$), that endpoint is declared statistically non-significant, further testing of these secondary endpoints is terminated, and no formal statistical conclusion is reached on the remaining endpoints, irrespective of their observed p-values.
Subgroup analysis of the secondary endpoints will be performed for the following subgroups: sex, age (Age<65, Age≥65), race (white vs. non-white), baseline UUI episodes, site, and region (US vs. OUS). The subgroup analyses of the secondary efficacy endpoints are regarded as exploratory and will not be used for labeling purposes and will be summarized with descriptive statistics.

The sponsor intends to file a PMA submission at the end of the six-months of follow-up if the null hypothesis associated with the primary efficacy endpoint is rejected. This will include analysis of the primary endpoint and the first secondary endpoint in the hierarchical sequence. If the analysis of the first secondary endpoint is statistically significant, the remaining secondary endpoints in the sequence will be analysed in order when the data are available.

5.6.3.2 Secondary Endpoints not Intended for Labeling Claims

Analysis of secondary endpoints not intended for labeling claims, (i.e., those secondary endpoints not included in the hierarchical testing in Section 5.6.3.1) will be considered exploratory in nature. Each of these endpoints will be tested at the $\alpha = 0.05$ level using a two-sided test with no adjustment for multiplicity. Any p-values from these analyses will not be reported in the labeling.

5.6.4 Handling of Missing Data

All attempts will be made to limit the amount of missing data. No attempt will be made to impute missing data for any analyses except for the primary and secondary efficacy endpoints as specified below in the mITT analysis set. If a data point is missing, that data point will not contribute to that portion of the analysis. The number of evaluable observations will be reported in the analysis so that extent of missing data can be assessed.

For the following efficacy endpoints, subjects withdrawing for reasons related to a lack of efficacy or due to an adverse event (including death) will be counted as failures if they do not have evaluable data for the endpoint:

- The percentage of all implanted subjects who experience an improvement in UUI symptoms of 50% or more at 6-month post-implant (primary efficacy endpoint)
- The percentage of all implanted subjects who experience an improvement in UUI symptoms of 50% or more at 12-month post-implant (secondary efficacy endpoint)

For the continuous efficacy endpoints that may be used for labeling claims, no change from baseline values will be assumed for subjects withdrawing for reasons related to a lack of efficacy or due to an adverse event (including death) if they do not have evaluable data for the endpoint.

If data on these endpoints are missing for other reasons (missing diary or ICIQ-OAB questionnaire, subject loss-to-follow-up, etc.), multiple imputation will be used to impute the outcome. Results from the imputed data sets will be combined appropriately to preserve valid inference. A tipping point analysis will also be performed, examining outcomes for the primary efficacy endpoint for all possible permutations of imputed values for missing subjects.

Multiple imputation will be based on the following variables and will use full conditional specification and 100 imputed data sets. Variables to be included in the imputation model include:

- baseline age, investigational site, baseline UUI episodes, and all available follow-up visit UUI episodes through 6 months or 12 months (for the primary and secondary efficacy endpoints of improvement in UUI)
- baseline age, investigational site, baseline score, and all available follow-up visit scores through 6 months or 12 months (for the remaining secondary efficacy endpoints that may be used for labeling claims)

If there are fitting issues due to the separation or quasi-separation of data points, an augmented likelihood approach will be used.

5.7 Analysis of Replacements, Explants

During the course of the study, it may be necessary for the investigational device to be replaced or permanently explanted from the subject. Revision rates in the literature range from 4% in the InSite study[ii] to 33% in the Herbison[vi] study. Depending on the timing of replacements, it may be necessary for the subject's response rate to be reassessed in the first 28 to 42-days post-implant.

During Phase I of the study, if a subject receives a neurostimulator replacement prior to Month 3 clinic visit, the subject will restart their follow-up schedule, and repeat the trialing period and any of the follow-up visits completed prior to replacement. If a subject receives a device replacement prior to the assessment of the primary efficacy endpoint in Phase II (Month 6 clinic visit), the subject will restart their entire follow-up schedule post-implant, including the responder rate assessments at Day 28 and Day 42, if not a responder at Day 28. The date of replacement will be considered the new Device Implant date for analysis and follow-up purposes. If a device replacement is required after the follow-up periods indicated above, subjects will continue on their original follow-up schedule and do not repeat any of the assessments or follow-up visits already completed.

If assessment of the response rate is repeated in the first 28 to 42-days post-implant, data collected after the most recent implant will be used in the analysis. Similarly, if the subject is required to repeat any portion of the post-implant follow-up visits after a replacement (replacement occurred before Month 6 clinic visit), the data from the replacement visit(s) will be used in analysis. The date of replacement will be considered the Device Implant date for analysis and follow-up purposes. A subject is only allowed to repeat any portion of the post-implant period once.

If the device is permanently explanted prior to the assessment of the primary efficacy endpoint for any reason, the subject will be considered a failure for the primary efficacy analysis.

5.8 Poolability Analyses

All investigational sites will follow the requirements of a common protocol and standardized data collection procedures and forms. The primary efficacy and safety endpoints will be presented separately for each site using descriptive statistics. Poolability of the primary efficacy endpoint across investigational sites will be evaluated using a logistic regression model with a fixed effect for site. If there is a significant effect of site (using a significance level of 0.15), the analysis will be repeated using

site as a random effect. Sites enrolling fewer than 5 subjects will be removed from these poolability analyses.

Additional exploratory analyses may be performed to understand any variations in outcome by site. Assessment of poolability by geography (US vs. OUS) will be performed based on logistic regression.

5.9 Safety Analyses

Investigators are required to report all AEs during the course of the study. Investigators are responsible for providing a description of all AEs, including the date of the event, treatment and the clinical outcomes for the subject. Investigators will evaluate each event for seriousness, severity and relatedness to the procedure, device or stimulation therapy. All adverse events, including unanticipated adverse device effects (UADE), and serious adverse events (SAE) will be reported for all enrolled subjects.

For Phase I, all adverse events reported over the course of the study will be summarized for all enrolled subjects. Any adverse events reported for consented subjects who were not enrolled will be reported separately.

For Phase II, all adverse events reported over the course of the study will be summarized for the mITT analysis set. Any adverse events reported for enrolled (i.e., consented) subjects who are excluded from mITT analysis set will be reported separately.

Adverse events will be tabulated with the number of events and subjects with event for each event type and overall. Rates will be reported as the number of subjects who experience at least one event during the analysis interval out of the total number of subjects with follow-up to the beginning of the analysis interval. Serious adverse events will also be tabulated. The rate of all AEs and SAEs reported in the study will be reported.

All AEs and SAEs will also be summarized by relatedness. Adverse events leading to death or study discontinuation will be provided in a listing format.

5.10 Interim Analyses

The sponsor will use Phase I, feasibility phase, of the trial to assess system performance, subject compliance and determination of the daily stimulation duration for Phase II. Phase I data will also be examined as needed to determine if modifications need to be made to training of the centers on implant technique, subject instructions, duration of stimulation, etc. Formal statistical conclusions from Phase I will not be made so there are not type I error concerns.

For Phase II, there are no planned interim analyses for the purpose of adaptive study modifications or early stopping for success. An interim analysis will be performed when the 60th Phase II subject has completed the 3-month post implant visit. For this interim analysis, the primary and secondary safety and effectiveness endpoints will be assessed, but at the three-month timepoint. This interim analysis will be used for business purposes only, no modifications to the study will be made based on this analysis, and outcomes will not be published or shared with investigators, patients, or Neuspera employees with direct subject access.

In order to ensure that adequate safety information is available to the agency, the sponsor commits both to completing the study to the original planned sample size of 120 attempted implants and to submitting updated information with additional follow-up safety data at the time of the 100-day meeting. Operational bias will be minimized by having a third party, independent statistician separate from study operational personnel and study statistician, who will perform the interim analysis.

5.11 Protocol Deviations and Device Deficiencies

Deviations from the procedures outlined in the protocol will be reported by investigational sites on case report forms. Protocol deviations will be summarized for all deviations and by type with event counts and number of subjects with at least one deviation.

Device deficiencies will be summarized in a similar fashion to protocol deviations.

6 Changes from Planned Analyses

Any changes to planned statistical analyses determined necessary prior to performing the analyses will be documented in an amended Statistical Analysis Plan and approved prior to the analysis when possible. Any other deviations or changes from the planned analyses deemed necessary due to violation of critical underlying statistical assumptions, data characteristics, or missing data will be clearly described in the clinical study report with justification and rationale.

7 Subject Listings

Subject listings will be provided for the primary endpoints. Listings will also be provided for adverse events.

8 Stopping Rules for Phase II

Phase II will use stopping rules designed to suspend enrollment and trigger a series of events if at any time there is sufficient statistical evidence to conclude that any of the Phase II probabilities of device

fracture, device dislodgement/migration, or surgical revision exceed comparator probabilities from an approved device. The comparator probabilities are shown in Table 2.

Table 2. Comparator event probabilities guiding the Phase II stopping rule. Values shown are point estimates (upper limits of 95% confidence intervals).

Event Type	Proportion of Subjects with Event
Fractures	0.014 (0.022)[1]
Dislodgements/Migrations	0.026 (0.036)[2]
Surgical Revisions	0.118 (0.239)[3]

Phase II will use the following stopping rules. At each occurrence of device fracture, dislodgement/migration, or surgical revision, form an exact one-sided binomial confidence interval for the event probability. Determine whether any of the following conditions is met:

- The lower limit of the interval for device fractures exceeds 0.014.
- The lower limit of the interval for device dislodgements/migrations exceeds 0.026.
- The lower limit of the interval for surgical revisions exceeds 0.118.

The numerator for each event is the number of subjects who have experienced the event up to that time. The denominator will be the number of subjects implanted up to that time. For device fractures and migrations/dislodgements, the confidence level of the interval will be 95%. For surgical revisions, the confidence level will be 80% if 50 or fewer subjects have been implanted; if more than 50 subjects have been implanted, the confidence level will be 95%. If any of the above conditions is met, the following actions will be taken:

- Both enrollment and implants will be suspended and FDA will be notified within 7 days
- Neuspera will conduct a root cause analysis to determine the cause of the safety events that triggered the stopping rule
- Neuspera will prepare a safety report for submission to the DSMB

[1] Based on Medtronic's [2]020 Product Performance Report showing 19 subjects experienced lead fractures out of 1,314 subjects followed (page 167).
[2] Based on Medtronic's 2020 Product Performance Report showing 34 subjects experienced lead dislodgements/migrations out of 1,314 subjects followed (page 167).
[3] Based on Medtronic Clinical Summary Report stating 6 out of 51 subjects had "surgical procedures to resolve adverse events or technical observations, or due to lack of efficacy … from the time of test lead implant to 6-month follow-up after full system implant" (page 16).

- The DSMB will convene, review the safety report and root cause analysis, and make a recommendation on study continuance (continue enrollment post FDA review and approval, continue enrollment suspension pending FDA decision and/or further investigation by the sponsor, etc.)
- Neuspera will submit to FDA the safety report, the root cause analysis, and the DSMB's recommendation
- FDA will make final determination on study continuance

The stopping rules will take effect only when more than 15 subjects have been implanted. After more than 15 subjects have been implanted, all events, including any events within the first 15 subjects, will be counted in assessing the criteria for stopping. However, if there are 4 or more surgical revisions when 15 or fewer subjects have been implanted, the stopping rule is triggered.

The use of an 80% confidence level as compared to 95% early in the study for surgical revisions makes it more, not less, likely that enrollment would stop due to surgical revisions in the event that revision rates are unacceptably high.

9 References

[i] McCrery R, Lane F, Benson K, et al. Treatment of Urinary Urgency Incontinence Using a Rechargeable SNM System: 6-Month Results of the ARTISAN-SNM Study. *J Urol*. 2020;203(1):185-192. doi:10.1097/JU.0000000000000458

[ii] Siegel S, Noblett K, Mangel J, et al. Results of a prospective, randomized, multicenter study evaluating sacral neuromodulation with InterStim therapy compared to standard medical therapy at 6-months in subjects with mild symptoms of overactive bladder. Neurourol Urodyn. 2015;34(3):224-230. doi:10.1002/nau.22544

[iii] Noblett K, Siegel S, Mangel J, et al. Results of a prospective, multicenter study evaluating quality of life, safety, and efficacy of sacral neuromodulation at twelve months in subjects with symptoms of overactive bladder. Neurourol Urodyn. 2016;35(2):246-251. doi:10.1002/nau.22707

[iv] Coyne K, Revicki D, Hunt T, Corey R, Stewart W, Bentkover J, Kurth H, Abrams P. Psychometric validation of an overactive bladder symptom and health-related quality of life questionnaire: the OAB-q. Qual Life Res. 2002 Sep;11(6):563-74.

[v] Coyne, K., et al., Determining the importance of change in the overactive bladder questionnaire. Journal of Urology. 2006: Aug; 176(2):627-32.

[vi] Herbison GP, Arnold EP. Sacral neuromodulation with implanted devices for urinary storage and voiding dysfunction in adults. Cochrane database Syst Rev. 2009;(2):CD004202. doi:10.1002/14651858.CD004202.pub2