Feasibility Study on Laser Interstitial Thermal Therapy Ablation for the Treatment of Medically Refractory Epilepsy (FLARE)

STATISTICAL ANALYSIS PLAN

MON-CL-1010, Rev. A

Prepared by:

[redacted]

Senior Principal Medical Research Biostatistician
NAMSA, Inc.

11 AUG 2017
NCT02820740

Table of Contents

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

1 Purpose

The purpose of this Statistical Analysis Plan (SAP) is to document analyses that are planned for the FLARE study before they are performed (*pre-specified analyses*). The results of the analyses documented here are to be presented in the clinical study final report. Selected analyses may also be presented in interim reports, DSMB reports, and manuscripts reporting study results, as deemed appropriate by the authors. Additional analyses of the study data beyond the analyses pre-specified in this plan are expected, therefore the Statistical Analysis Plan does not preclude ad hoc analyses that may provide additional useful description of the performance of the investigational device.

2 Scope

This SAP should be read in conjunction with the Clinical Study Protocol (CL10054) and Case Report Forms (MON-CRF-1001 through MON-CRF-1030). Changes to the versions of the Clinical Study Protocol (protocol) and Case Report Forms (CRFs) cited above may necessitate updates to the SAP.

3 Applicable Documents

Document Number	Document Title
CL10054	Monteris FLARE Clinical Study Protocol, Rev A
MON-CRF-1001 through MON-CRF-1030	Monteris FLARE Study Case Report Forms
NAMSA STATSOP-002	Statistics Standard Operating Procedure – Statistical Analysis Plan

4 Software

Statistical analyses, tables, listings, and figures will be primarily produced using SAS statistical software (SAS Institute, Cary, NC), version 9.2 or later, or another validated statistical software package.

5 Abbreviations

The following is a list of abbreviations used in the body of this document.

ADE	adverse device effect
AE	adverse event
AED	antiepileptic drug
BVMT-R	Brief Visual Memory Test – Revised
BNT-60	Boston Naming Test – 60 item version
CEC	Clinical Events Committee
DSMB	Data Safety and Monitoring Board
CRF	case report form
ICF	informed consent form
ICH	International Conference on Harmonization
IDE	Investigational Device Exemption
IQ	intelligence quotient
ILAE	International League Against Epilepsy
LITT	laser interstitial thermal therapy
MRI	magnetic resonance imaging

MTLE	mesial temporal lobe epilepsy
NBS	NeuroBlate® System
QOLIE-31	Quality of Life in Epilepsy – 31 item version
RAVLT	Rey Auditory Verbal Learning Test
SAE	serious adverse event
SOP	standard operating procedures
UADE	unanticipated adverse device effect
USADE	unanticipated serious adverse device effect
VAS	visual analog scale
WAIS-IV	Wechsler Adult Intelligence Scale-IV
WMS-IV	Wechsler Memory Scale-IV

6 Design

The FLARE study is a multicenter, open-label, prospective feasibility study to characterize the performance of laser interstitial thermal therapy (LITT) ablation using Monteris NeuroBlate® System (NBS) for the treatment of drug refractory mesial temporal lobe epilepsy in subjects who are candidates for surgical resection.

The study design is outlined in Figure 1.

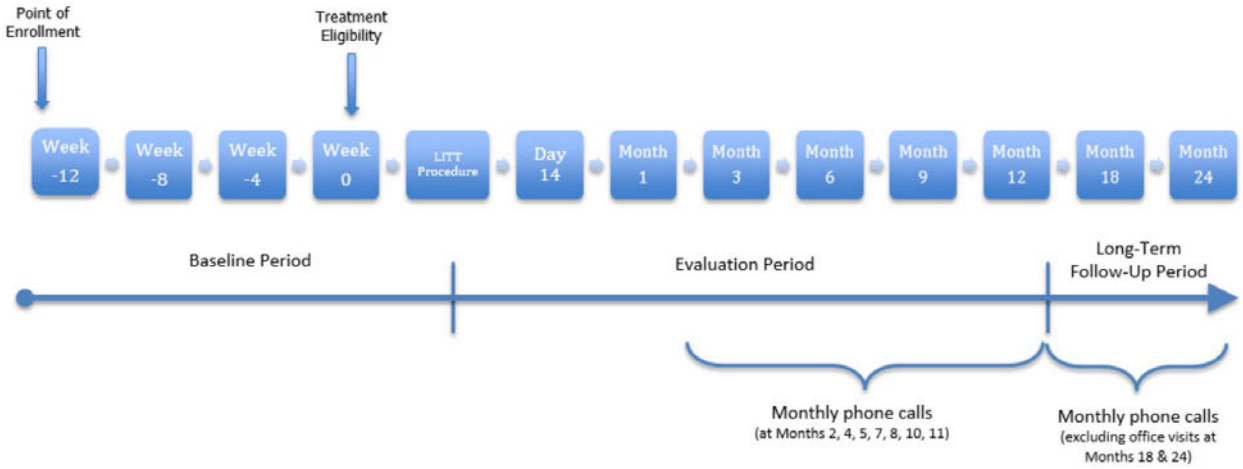

Figure 1: Study Design Schema

7 Endpoints

The primary endpoint for the study is:

- To characterize safety (adverse events and neuropsychological changes) of LITT for the treatment of drug-refractory medial temporal lobe epilepsy

The secondary endpoints for the study are:

- To characterize seizure outcome (based on seizure frequency and surgical outcome classifications) of study subjects treated with LITT
- To characterize the quality of life of study subjects treated with LITT

8 Data Collection

Table 1 outlines the visits, visit windows, and testing requirements at each evaluation time point from Baseline through the Month 12 visit. Table 2 outlines the visits, visit windows, and testing requirements at each evaluation time point from the Month 13 visit through the Month 24 visit.

The remainder of the page is intentionally blank.

Table 1: Visits and Testing Requirements and Data Collection Summary (Baseline through Month 12)

	Pre-Procedure						Post-Procedure												
	Week -12	Week -8	Week -4	Week 0	LITT Proc.	DC	Day 14	Mo. 1	Mo. 2	Mo. 3	Mo. 4	Mo. 5	Mo. 6	Mo. 7	Mo. 8	Mo. 9	Mo. 10	Mo. 11	Mo. 12
Visit Window	*Within 7 days of enrollment*	*± 7 days*	*± 7 days*	*± 7 days*	*Within 30 days of Week 0*	*Day of DC*	*± 4 days*	*± 4 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*	*± 7 days*
Visit Type	🗣	🗣	🗣	🗣	🗣	🗣	🗣	🗣	☎	🗣	☎	☎	🗣	☎	☎	🗣	☎	☎	🗣
Inclusion and exclusion criteria	√																		
Informed consent	√																		
Demographics	√																		
Medical, surgical & epilepsy history	√																		
QOLIE-31	√												√						√
Medication review (epilepsy and non-epilepsy)	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Physical exam	√					√		√					√						√
Neurological exam	√					√		√		√			√			√			√
Formal visual field testing	√												√						
Serum pregnancy test (if female of child-bearing potential)	√			√															
Neuropsychological testing [1]	√[2]		√[2]							√			√						√
Functional assessment	√			√			√	√		√			√			√			√
Seizure classification assessment	I									U			U						U
Seizure diary	D	C/D	C/D	C		D	C/D	C/D	C	C/D	C	C	C/D	C	C	C/D	C	C	C/D
Adverse event monitoring		√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Treatment criteria				√															
LITT Treatment					√														
MRI					√	√							√						
Surgical pain VAS						√	√												
Surgical outcome classifications[3]										√			√			√			√
Subject satisfaction													√						√

Abbreviations: C; collect; D, dispense; DC, discharge; I, initial; LITT, laser interstitial thermal therapy; Mo., month; MRI, magnetic resonance imaging; QOLIE-31, Quality of Life in Epilepsy - 31; U, update; VAS, visual analog scale.

[1] See Appendix A for details of the neuropsychological testing at each visit. These can be performed on a separate day than the study visit, as long as it is completed within the study window.

[2] Prior neuropsychological testing may be used for baseline values if the testing was conducted within 6 months prior to the specific baseline visit and there were no significant changes to the subject's AEDs or changes to the subject's clinical condition that may affect cognitive function.

[3] See Appendix B for details about the Engel and ILAE surgical outcome classifications.

Table 2: Visits and Testing Requirements and Data Collection Summary (Month 13 through Month 24)

	Months 13, 14, 15, 16, & 17	Month 18	Months 19, 20, 21, 22, & 23	Month 24	Early Withdrawal
Visit Window	*± 7 days*	*± 14 days*	*± 7 days*	*± 14 days*	
Visit Type	☎	🏠	☎	🏠	
QOLIE-31		√		√	√
Medication review (epilepsy and non-epilepsy)	√	√	√	√	
Neurological exam		√		√	√
Functional assessment		√		√	√
Seizure diary	C/D	C/D	C/D	C	
Adverse event monitoring	√	√	√	√	√
Surgical outcome classifications[1]		√		√	√
Subject satisfaction		√		√	√
Neuropsychological testing[2]					√

Abbreviations: C; collect; D, dispense; QOLIE-31, Quality of Life in Epilepsy - 31; U, update.
[1] See Appendix B for details about the Engel and ILAE surgical outcome classifications.
[1] See Appendix A for details of the neuropsychological testing.

9 Sample Size

For this feasibility study, there are no hypothesis tests and therefore sample size is not based on power. Rather, the sample size of 30 subjects is based on providing a sufficiently robust characterization of adverse events and other study endpoints. With a sample size of 30 subjects, adverse events that have a 5% probability per-subject have approximately an 80% chance of being observed in at least one subject in the study. Study results obtained from a sample of 30 subjects will permit planning for the potential range of results that could occur in a larger pivotal trial.

Up to 45 subjects will be enrolled at up to 8 sites in the United States to ensure that approximately 30 subjects undergo the LITT procedure. It is expected that each site will enroll no more than 10 subjects to ensure that all sites have an opportunity to contribute data to the study and to prevent any one site from having an outsized influence on the overall results of the study.

10 Analysis Populations

The following analysis populations are defined:

- Enrolled Set: All subjects enrolled in the study
- Treated Set: All subjects that received thermal delivery with the NBS
- Per-protocol Set: A subset of the Treated Set that does not have a protocol deviation (see below)

In the feasibility study setting, conclusions about performance are best informed by analyses of the Treated Set. Analyses may be repeated on the Enrolled Set for full reporting of study results and on the Per-protocol Set to provide insight into the potential impact of protocol deviations on the analysis results.

The Per-protocol Set will exclude subjects with protocol deviations that seriously affect the integrity of the data. The deviations that will result in subject data being excluded from the Per-Protocol Set are:

1. Subject did not meet the following study inclusion criteria:

- Diagnosis of unilateral medial temporal lobe epilepsy (MTLE) confirmed clinically and with either (1) ictal scalp recording and MRI evidence of mesial temporal sclerosis or (2) intracranial ictal onset consistent with hippocampal origin.
- Subject or legally authorized representative is able to provide appropriate consent to participate.
- Subject averages 3 or more complex partial seizures (with or without secondary generalization) per month over the 3 most recent consecutive months prior to baseline.

2. Subject should not have undergone the study procedure (i.e. subject was found not to have met the treatment eligibility criteria in Section 6.4 of the clinical study protocol)

Protocol deviations may be reviewed prior to performing analysis of study outcomes to identify deviations, in addition to those identified above, that may result in excluding a subject from the per-protocol set. Reviewing protocol deviations prior to analysis of outcomes is intended to prevent bias in the analysis results while still allowing for identification of significant protocol deviations that could impact outcomes but that were not anticipated at the time that the Statistical Analysis Plan was written.

The reasons that Treated Set subjects are excluded from the Per-protocol Set will be tabulated.

11 Statistical Analyses

11.1 Endpoint Analysis

Since the objective of this feasibility study is to characterize the performance of LITT, no formal hypotheses tests are planned for any endpoint. Descriptive statistics and exploratory analyses will be used in reporting outcomes for all endpoints.

11.1.1 Primary Endpoint: Analysis of Adverse Events

Adverse events will be presented in summary tables displaying the number of events, the number of enrolled subjects with ≥ 1 event, and the proportion of enrolled subjects with ≥ 1 event. The rates of adverse events will serve as the primary analysis.

Adverse event analysis may be repeated for only those events that occur after the initiation of treatment with rates calculated as a proportion of treated subjects..

Phase

The timing of adverse events relative to the initial LITT procedure will be tabulated by reporting the event counts, subject counts, and subject rates of events that occur within the following phases:

- **Baseline:** Prior to LITT procedure
- **Operative:** From the beginning of the LITT procedure through 30 days post-procedure
- **Year 1:** From 31 days post-procedure through 12 study months (12 month visit date or close of 12 month visit window if 12 month visit is missed)
- **Year 2:** After 12 study months through 24 study months (24 month visit date or close of 24 month visit window if 24 month visit is missed)

Additional analyses may be performed that divide the above phases into more detailed time intervals or that combine phases to present an overall summary of all adverse events through a specific time point (e.g. Operative Through Year 2).

Severity

The severity of adverse events will be tabulated by reporting the event counts, subject counts, and subject rates of events that are:

- **Mild:** Does not interfere in a significant manner with the subject's normal functioning level.
- **Moderate:** Produces some impairment of functioning but is not hazardous to health.
- **Severe:** Produces significant impairment of function or incapacitation and is a hazard to the subject's health.

Relationship

The relationship of adverse events to the device treatment or procedure will be tabulated by reporting the event counts, subject counts, and subject rates of events based on the following relationship categories:

- **Definite** – The adverse event follows a strong temporal sequence to the receipt (or attempted receipt) of the device treatment or procedure. This can include an adverse event that occurs after the study procedure.
- **Probable** – The adverse event follows a reasonable temporal sequence to the receipt (or attempted receipt) of the device treatment or procedure, and the possibilities of other factors, such as underlying and concomitant illness, concomitant medications, or concurrent treatment can be excluded.
- **Possible** – The adverse event follows a reasonable temporal sequence from receipt of the device treatment or procedure and the possibility of device treatment or procedure involvement cannot be excluded. However, other factors such as underlying or concomitant illness, concomitant medications, or concurrent treatment are presumable.
- **Unlikely** – The adverse event has an improbable temporal sequence to the receipt of the device treatment or procedure, or it can be reasonably explained by other factors, including underlying or concomitant illness, concomitant medications, or concurrent treatment.
- **Not Related** – The adverse event has no temporal sequence to the LITT procedure, NeuroBlate System or any user handling, or it can be explained by other factors, including underlying disease or concomitant illness, concomitant medication, or concurrent treatment.

In addition, the relationship may be characterized in a binary way as Related or Unrelated where Related corresponds to the categories Definite, Probable, Possible and Unrelated corresponds to the categories Unlikely, Not Related.

For events that are Unrelated to the device treatment or procedure, the relationship to the following factors will be tabulated:

- **New illness or injury:** Event is clearly attributable to a new illness or injury with no temporal relationship to the device, procedure, treatment or a medication.
- **Pre-existing condition/disease:** Event is attributable to an underlying condition/disease with no temporal relationship to the device, procedure, treatment or medication.
- **Epilepsy medication:** Event has a strong temporal relationship to an epilepsy medication.
- **Medication (non-epilepsy):** Event has a strong temporal relationship to a non-epilepsy medication.
- **Anesthesia**: Event has a strong temporal relationship to anesthesia.
- **Other**: Event has a relationship to another potential causation
- **Unknown:** Event relationship is not known or unsure.

Sets of Events

The tabulation of event type and cross-tabulation of event type by severity and relationship will be repeated, at a minimum, for the following sets of events:

- All Adverse Events (AEs)
- Serious Adverse Events (SAEs)
- Adverse Device Effects (i.e. AEs related to the use of the study device. This will include any event that is assessed as definitely, probably or possibly related to the device or the LITT procedure)
- Serious Adverse Device Effects (SADEs)
- Unanticipated Adverse Device Effect (UADE) / Unanticipated Serious Adverse Device Effect (USADE)

If any of the sets of events above does not occur in the study, that will be noted and no tabulation is required for that set.

11.1.2 Primary Endpoint: Analysis of Neuropsychological Changes

For each neuropsychological test, descriptive statistics will be reported for the measured score at baseline and each follow-up visit. In addition, change from baseline will be calculated at each follow-up visit and descriptive statistics will be reported. For each subject, the change from baseline will be categorized as improved, unchanged, or deteriorated based on whether the change exceeds a pre-specified threshold determined by Reliable Change Index (RCI) methodology for that specific neuropsychological test with a 95% confidence level. The threshold thus represents a degree of change that is unlikely (<2.5% chance in either direction) to occur due to measurement error of the neuropsychological instrument alone. The proportion of subjects in each category at each follow-up time will be tabulated.

For each neuropsychological test, the RCI with 95% confidence level is calculated as 1.96 x S_{diff} where S_{diff} is the standard error of the difference between the baseline and follow-up neuropsychological test score. The standard error of measurement is used to calculate the standard error of the difference.

$$S_{diff} = \sqrt{SEM_1^2 + SEM_2^2}$$

The standard error of measurement is calculated based on the standard deviation of the test score across subjects (SD) and the test-retest reliability (r) of the test with the equation

$$SEM_1 = SD_1\sqrt{1 - r}$$

In order to approximate S_{diff} for the population of subjects enrolled in the study, the observed standard deviation at baseline (SD1) will be used to calculate both SEM1 and SEM2. The approach assumes that the variation SD2 would be the same as SD1 if all subjects were re-tested at the specified follow-up interval without having undergone treatment with the study device. Since SD2 cannot be observed, it is approximated by SD1. The test-retest reliability for each test is obtained from published literature and will be documented prior to performing analysis.
A particular subject's change from baseline in a neuropsychological test score will be considered statistically different from measurement error if the observed difference exceeds 1.96 x Sdiff.

As there is no control group for comparison, any observed changes in neuropsychological function will need to be evaluated in comparison to longitudinal changes observed in prior longitudinal studies of similar subjects, examining typical changes in both subjects that have and have not undergone invasive therapy. In addition, since up to 2.5% of subjects are expected to have a change as large as the RCI in each direction (positive and negative), observed proportions will be interpreted based on the degree to which they differs from 2.5%.

For the primary analysis of neuropsychological changes, English and Spanish language versions of the same test will be considered interchangeable and pooled together for analysis. Sensitivity analyses will be performed by analyzing English and Spanish versions separately.

11.1.3 Secondary Endpoint: Analysis of Seizure Outcomes

Surgical outcome classification will be reported by tabulating the Engel and ILAE class frequency distribution at each scheduled follow-up, among subjects completing that follow-up visit. The change in surgical classification between visits may be described by performing a cross-tabulation of the classifications at subsequent visits. The classifications are described in more detail in Appendix B.

For the Engle classification, there are several classification terms that can be subjective in nature, so further definitions are provided to the sites, as follows:

- the term "disabling seizures" includes all focal seizures *except*
    - those without impairment of consciousness or responsiveness (simple partial seizures) involving subjective sensory or psychic phenomena only
- The term "worthwhile improvement" is defined as ≥50% seizure reduction compared to baseline.

Seizure frequency will be characterized by calculating the rate of seizure occurrence within specific follow-up time periods (count of reported events/reported time frame). Follow-up time periods of interest include all follow-up and between-visit follow-up periods (e.g., procedure to Month 3) to characterize the longitudinal course of seizure frequency. For each time period, descriptive statistics of subject-specific rates characterize the distribution of observed seizure rates. The change from baseline in seizure rates may also be calculated by calculating the difference in rate between a follow-up period and the baseline rate of seizures occurring prior to treatment (count of reported events/reported time frame). Since subjects may have different baseline seizure rates, the relative change from baseline in the event rate (r)

$$\frac{(r_{follow-up} - r_{baseline})}{r_{baseline}}$$

will also be calculated to scale the degree of improvement relative to the baseline rate. These analyses may be repeated for subsets of seizures with specific classifications to characterize the frequency of different seizure types.

The primary analysis of seizure frequency will include subjects with at least 70 completed diary days in each 3 month period from procedure through 12 months and at least 140 diary days in each 6 month period from 12 months through study completion. To account for incomplete diary completion, sensitivity analyses may be performed by including subjects with less than the required level of diary completion and by using the number of days with completed diary entries to determine the denominator in the event rate calculation.

Seizure frequency may be further characterized by calculating the *proportion of responders*. A responder is a subject whose seizure frequency in a specified 3-month interval is reduced by ≥50% as compared to the 3-month Baseline Phase.

Seizure free days may be further characterized by calculating the percent change in seizure-free days that subjects experience during a specified 3-month interval compared to Baseline. In order to make the 3 month follow-up intervals and the Baseline Phase comparable, the number of days actually recorded will be prorated to an 84-day interval for both intervals. Seizure-free days may also be expressed as a proportion of the number of follow-up days with available diary data.

Freedom from seizures with altered awareness will be reported using Kaplan-Meier curve analysis. Seizures with altered awareness include those categorized as either “impairment of consciousness or responsiveness” or “evolving to a bilateral, convulsive seizure”. The time to event analysis will exclude seizures that occur within the first 4 weeks post-procedure. The event time will be calculated using the earliest post-treatment date of a seizure with altered awareness, as reported in subject diaries. The censoring date will be the last completed diary date for subjects without seizures with altered awareness. The Kaplan-Meier analysis may be repeated for freedom from all seizures, seizures with other classifications, and inclusion of seizures within the first 4 weeks post-procedure. The log-log transformation will be used to calculate 95% confidence intervals. The time point of primary interest is 24 months.

Longest seizure-free interval will be characterized with descriptive statistics for each 3-month period and cumulatively over consecutive 3-month periods. The seizure-free interval will be

measured as the number of completed diary days without a seizure since the most recent post-treatment seizure or the beginning of the time period being characterized (e.g. since the day of treatment for the cumulative analysis), whichever is more recent. The longest seizure-free interval will also be calculated for the baseline period, which permits calculation of the change in the longest seizure-free interval from baseline for each subject. The change in longest seizure-free interval may be characterized with descriptive statistics.

Rescue medication use will be analyzed with both a Kaplan-Meier analysis to describe the freedom from use of any rescue medication and a descriptive statistics analysis of the frequency of rescue medication use (number of occasions that rescue medication is required / follow-up time). These analyses will be performed using the same methods as the freedom from seizures and seizure frequency analyses above, but with recue medication use as the event for analysis rather than seizures.

11.1.4 Secondary Endpoint: Analysis of Quality of Life

Quality of life is assessed with the QOLIE-31 instrument. QOLIE-31 scores at baseline and each scheduled follow-up visit will be characterized using descriptive statistics. In addition, the per-subject paired change from baseline will be calculated and descriptive statistics reported for each scheduled follow-up visit. These analyses may be repeated for the QOLIE-31 sub-domains.

Longitudinal analysis of the QOLIE-31 scores may be performed using a GEE repeated measures model with a working AR1 correlation structure to estimate the mean (with 95% confidence intervals) of the QOLIE-31 over time while accounting for within subject correlation. Baseline QOLIE-31 and visit (categorical) will be included as covariates in the model. The estimate of QOLIE-31 score at 24 months is of primary interest.

Finally, the proportion of subjects with a change from baseline in QOLIE-31 score of 13 points or more will be calculated at each follow-up visit. The threshold of 13 points is based on previous literature estimating 13 points of QOLIE-31 to be the minimum clinically important difference.[1]

Further analyses related to subjects' quality of life will be performed with descriptive statistics of responses to the functional assessment (employment status, education level, driving status, and living arrangements) and subject satisfaction with the results of treatment.

12 General Statistical Considerations

12.1 Descriptive Statistics

Standard summary statistics will be calculated for all study variables to be reported. For continuous variables, statistics will include means, standard deviations, medians and ranges. Categorical variables will be summarized by frequency distributions.

Confidence intervals may be reported, as appropriate, using the methods listed below. A 95% confidence interval for the mean may be reported using the t-distribution method. A 95% confidence interval for the proportion may be reported using the Clopper-Pearson exact method.

[1] Wiebe S, Matijevic S, Eliasziw M, Derry PA. Clinically important change in quality of life in epilepsy. J Neurol Neurosurg Psychiatr 2002;73(2):116-20.

12.2 Disposition of Subjects

All subjects that sign the informed consent form (ICF) will be accounted for by reporting

1. Number of subjects that sign the ICF/enroll
2. Number of subjects that enroll and meet all eligibility criteria
3. Number of subjects that are eligible for treatment (complete the Baseline Phase and meet the treatment criteria)
4. Number of subjects that exit the study prior to LITT treatment
5. Number of subjects that undergo LITT treatment
6. Number of subjects that complete each follow-up visit
7. Number of subjects that exit the study after LITT treatment but prior to the Month 24 visit.

12.3 Duration variables

Months for this study are defined as 28-day intervals and are referred to as “study months” to distinguish the interval from calendar months. Duration in study months is calculated is the number of days divided by 28. Duration in study years is calculated as the number of days divided by 12 study months.

12.4 Rate variables

Rates will be calculated as the ratio of the number of events divided by the number of days over which data collection occurred. Rates may be reported as events per day, per study month, or per study year. For diary data, days on which episode data is not available will not be included in the denominator.

12.5 Protocol deviations

The number of protocol deviations, number of subjects in which the deviations occur, and proportion of subjects with at least one deviation will be reported for all deviations and by sub-categories of deviations (e.g. informed consent, visit compliance, etc.).

12.6 P-values

While no hypothesis tests are specified, p-values may be used descriptively to report comparisons among sub-groups of subjects or to describe one-sample changes from baseline. P-values will be reported to a maximum of three decimal places with values < 0.001 being denoted as “<0.001”. For p-values greater or equal to 0.01, two decimal places will generally be reported. One-sided p-values < 0.025 and two-sided p-values less than 0.05 will be considered statistically significant.

12.7 Time Point of Interest

The study was designed to follow subjects through 24 months post-procedure and 24 months has been identified as a time point of primary interest. Analyses will be performed to assess the consistency of study endpoints over the 24 month time period and the degree of predictability of later outcomes based on outcomes observed at earlier time points (e.g. correlation between seizure rate for the time period of 6 to 12 months and seizure rate for the time period of 12 to 24 months).

13 Subgroup Analyses

Descriptive subgroup analyses for each endpoint will be performed to characterize the consistency of overall study estimates across the following pre-specified groups:

1. Age (tertiles of age)
2. Race (white vs non-white)
3. Gender (male vs female)
4. Baseline seizure frequency (tertiles of baseline seizure frequency)

5. Time since epilepsy diagnosis (tertiles of years since diagnosis)
6. Prior/concurrent VNS treatment (yes vs no)

Subgroup results will only be reported for subgroups that have at least 5 subjects since the smaller the sample size in a subgroup, the less likely it is that the observed results will be generalizable to the subgroup population.

Additional ad hoc subgroup analyses may be performed after all subjects have been enrolled, depending on the subject demographics and available sample size.

14 Treatment Factors Analyses

Descriptive analyses for each endpoint will be performed to characterize the potential relationship between treatment factors and study endpoints.

Treatment factors that will be analyzed include:

1. Complete versus incomplete LITT procedure
2. Abnormal findings on post-op images (e.g. edema, hemorrhage, infarction, or collateral damage)
3. Trajectory anatomic measurements
4. Ablation volume

15 Data Handling

15.1 Partial Dates

In the case that only partial dates are available for adverse events, the dates will be imputed. If only the month and year are known, any duration calculated with the partially known date will assume that the date is the first day of the known month. If only the year is known, any duration calculated with the partially known date will assume that the date is July 1 of the known year. In the event that the imputed date is before the procedure date, the imputed date will be equal to the procedure date. If the imputed date is after the last follow-up date or event resolution date, the imputed date will be the latest possible date.

15.2 Visit Windows

All data attributed to a time point per the CRF will be included in the analysis of that time point, regardless of whether the actual visit date was out of window.

15.3 CEC Adjudication

In all tabulations of adverse events, the categorizations of adverse events determined by the CEC (e.g. relatedness of adverse events to the device treatment or procedure) will be reported, if available, and categorizations provided by investigators will be reported otherwise. Adverse event listings will report both CEC and investigator categorizations.

16 Interim Analyses

Analyses will be reported on data from baseline through a major study time point once all subjects have completed that major study time point (e.g., 6 months, 12 months and 24 months). There is minimal opportunity for bias since all procedures will have been completed before any analysis is reported and all data will have been collected for a time point before analysis of that time point is reported.

The study's DSMB may request descriptive analyses of accumulating data at its discretion.

17 Missing Data Analyses

Every effort will be made to reduce the incidence of missing data. The study will be conducted with proper screening of study subjects, complete training of participating investigators, study coordinators and monitors. All subject data that are available on subjects who drop out during the study will be included.

The number and proportion of subjects eligible for and compliant with each follow-up examination will be presented. Subjects who withdraw from the study will be tabulated with reasons for withdrawal. Sensitivity analyses, such as multiple imputation or worst case imputation, may be performed to assess the robustness of study conclusions to the potential impact of missing data.

In worst case imputation, for binary endpoints the worse of the two possible outcomes will be imputed. The timing of the imputed event will be the last known event-free date for subjects that exit the study prematurely. For continuous endpoints assessed at a single time period (e.g. quality of life), the worst observed value among subjects with observed data at that time period will be imputed for subjects with missing data. For rates corresponding to a specific follow-up time interval, any partial observed data for a subject will be used. Data for the missing portion of the subject's follow-up time interval will be imputed using the worst observed value among all subjects with observed data during the corresponding missing portion of the follow-up time interval, in order to calculate an imputed rate for the overall follow-up time interval.

18 Sensitivity Analyses

Sensitivity analyses may be performed to describe the potential impact of various factors on the observed study results.

18.1 Medication and Concomitant Epilepsy Treatment

Changes to medications have the potential to impact the study endpoints and may be confounded with the impact of the investigational treatment. The analyses described below can provide insight into the potential impact of medication changes.

1. Subjects that had significant changes to the AED regimen (e.g., dosage change, discontinued AED, or new AED) between the Baseline Week -12 visit and the Month 24 visit will be excluded from analyses of visits after the AED changes occur (data censored after medication change).

Additionally, compare observed (uncensored) results at different follow-up time points between the 4 groups of subjects below. The AED regimen change group is determined by changes that occur prior to the follow-up time point being evaluated.

(a) No change in AED regimen (same medications and dosage compared to baseline)
(b) Increase in AED regimen (increased number of medications or dosage compared to baseline)
(c) Decrease in AED regimen (decreased number of medications or dosage compared to baseline)

(d) Other AED regimen change (increased dosage of a medication or addition of a new medication and decreased dosage or discontinuation of a different medication compared to baseline)

2. Subjects that had clinically significant changes to other concomitant epilepsy treatment (e.g., other surgical procedures, vagal nerve stimulation) between the Baseline Week -12 visit and the Month 24 visit will be excluded from analyses of visits after the changes to concomitant epilepsy treatment occur (data censored after concomitant epilepsy treatment change).

19 Poolability of Data

This study is designed and conducted as a multicenter clinical trial. All participating sites will be selected using the same criteria and trained prior to enrolling subjects. All subjects will be treated and evaluated following the same protocol to ensure generalizability of the study results. Poolability of the data across centers in a multicenter study is assumed without burden of proof.

Nevertheless, poolability of data across sites will be examined for apparent violation of the poolability assumption. The results of a poolability analysis will be used to assess the robustness of the study results, but not to change the study conclusion. For the primary outcomes, the rate of related adverse events and rate of changes exceeding the reliable change index for each core neuropsychological test will be performed separately for each participating institution. Additionally, a chi-square statistic will be used to describe the degree of variation in these proportions across sites.

For the seizure outcomes secondary objective, a Poisson regression will be performed with the post-procedure event rate over 24 months as the response variable and baseline event rate and site as explanatory factors, to assess variation in post-procedure seizure rates across sites.

For the quality of life secondary objective, an analysis of variance (ANOVA) will be used to assess homogeneity of the mean change in QOLIE-31 scores across sites at 24 months.

Sites with fewer than 5 subjects will be combined into a “small numbers” site for these analyses since the smaller the number of enrollments at a site, the less likely it is that the observed results reflect the expected results if the site had a greater number of enrollments. Results may also be presented separately for each site to describe the study endpoints across sites.

20 Deviations from the Statistical Analysis Plan

Any deviation from these planned statistical methods will be documented and discussed in the clinical study report along with statistical rationale for deviation.

APPENDIX A NEUROPSYCHOLOGICAL TESTING

This study will include a **core** neuropsychological battery as well as a **supplementary** battery, which are described in detail in the study protocol and listed in the tables below. The tests differ based on a subject's spoken language.

Table 3: Neuropsychological Testing Requirements for English-Speaking Subjects

Test	Baseline Week -12	Baseline Week -4	Month 3	Month 6	Month 12	Early Withdrawal
CORE						
WAIS-IV	X				X	X
RAVLT		X			X	X
BNT-60		X			X	X
COWAT		X			X	X
Animal Naming		X			X	X
Trail Making		X			X	X
Grooved Pegboard		X			X	X
Supplementary						
Logical Memory		Form 1	Form 2	Form 3	Form 1	
Auditory Naming		Form 1	Form 2	Form 3	Form 1	
Visual Naming		Form 1	Form 2	Form 3	Form 1	
Verbal Fluency		Form 1	Form 2	Form 3	Form 1	

Table 4: Neuropsychological Testing Requirements for Spanish-Speaking Subjects

Test	Baseline Week -12	Baseline Week -4	Month 3	Month 6	Month 12	Early Withdrawal
Core						
Raven's Matrices	X				X	X
WHO-UCLA AVLT		X			X	X
Ponton-Satz BNT		X			X	X
COWAT		X			X	X
Animal Naming		X			X	X
Color Trails		X			X	X
Grooved Pegs		X			X	X

APPENDIX B SURGICAL OUTCOME CLASSIFICATIONS

The two surgery outcome classifications used in this study, the Engel system,[2] and ILAE system,[3] are presented in Table 5 and Table 6.

The occurrence of early postoperative seizures or auras in the first week after epilepsy surgery is relatively common and may result from the effects of the acute surgical injury (termed "neighborhood" seizures). As such, both the Engel class I and the ILAE class 1 exclude "early" seizures. The Engel classification is not specific for the exact time period (first few weeks postoperative), whereas the ILAE classification excludes seizures in the first 4 weeks. *For the purposes of this study, seizures occurring in the first 4 weeks postoperative will be excluded when evaluating both the Engel and ILAE classes*.

An additional ILAE classification, 1a, is available for those that have been seizure- and aura-free since surgery (as recommended by Wieser et al).

Table 5: Engel Classification

Class I: Free of disabling seizures
A: completely seizure-free since surgery
B: non-disabling simple partial seizures only since surgery
C: some disabling seizures after surgery, bur free of disabling seizures for at least 2 years
Class II: Rare disabling seizures ("almost seizure-free")
A: initially free of disabling seizures but has rare seizures now
B: rare disabling seizures since surgery
C: more than rare disabling seizures since surgery, but rare seizures for the last 2 years
D: nocturnal seizures only
Class III: Worthwhile improvement
A: worthwhile seizure reduction
B: prolonged seizure-free intervals amounting to greater than half the follow-up period, but not < 2 years
Class IV: No worthwhile improvement
A: significant seizure reduction
B: no appreciable change
C: seizures worse

Table 6: ILAE Classification

Class 1: Completely seizure-free; no auras
Class 1a: *Completely seizure- and aura-free since surgery*
Class 2: Only auras[1]; no other seizures
Class 3: One to three seizure days[2] per year; ± auras
Class 4: Four seizure days per year to 50% reduction of baseline seizure days[3]; ± auras
Class 5: Less than 50% reduction of baseline seizures days to 100% increase of baseline seizure days; ± auras
Class 6: More than 100% increase in baseline seizure days; ± auras
[1] Auras are only counted if they are short in duration and similar or identical to the preoperative auras.
[2] A "seizure day" (for the purposed of this classification) is defined as a 24-hour period with one or more seizures. This may include an episode of status epilepticus.
[3] The number of "baseline seizure days" is calculated by determining the seizure-day frequency during the 12 months before surgery, with corrections for the effects of AED reduction during diagnostic evaluations prior to the index procedure.

[2] Engel J Jr, Van Ness PC, Rasmussen TB, Ojemann LM. Outcome with respect to epileptic seizures. In: Engel J Jr, editor. Surgical treatment of the epilepsies. New York: Raven Press; 1993:609-21.

[3] Wieser HG, Blume WT, Fish D, Goldensohn E, Hufnagel A, King D, et al. ILAE Commission Report. Proposal for a new classification of outcome with respect to epileptic seizures following epilepsy surgery. Epilepsia 2001;42:282–6.