STATISTICAL ANALYSIS PLAN

Latera RCT

Protocol Number: CP04

Protocol Version:	Version 1.0, 16 October 2017
Sponsor:	Spirox Inc.
Contract Research Organization:	Syntactx

STATISTICAL ANALYSIS PLAN

Approval Signatures

David Snead, PhD Date

Senior Principal Biostatistician, Syntactx

Jeremy Lao Date

Biostatistician, Syntactx

Elisa Hebb Date

VP, Clinical & Regulatory Affairs, Spirox Inc.

Revision History

Version Number	Version Date	Affected Section(s)	Summary of Revisions Made:
1.0	12-FEB-2018	Original	Initial Release

TABLE OF CONTENTS

Page

STATISTICAL ANALYSIS PLAN

STATISTICAL ANALYSIS PLAN

1 ABBREVIATIONS AND ACRONYMS

Acronym	Description
AE	Adverse Event
BMI	Body Mass Index
CI	Confidence Interval
CPAP	Continuous Positive Airway Pressure
DMC	Data Monitoring Committee
ESS	Epworth Sleepiness Scale
FESS	Functional Endoscopic Sinus Surgery
ICF	Informed Consent Form
IRB	Investigational Review Board
ITT	Intention-to-Treat
LWI	Lateral Wall Insufficiency
MedDRA	Medical Dictionary for Regulatory Activities
mITT	Modified Intention-to-Treat
NAO	Nasal Airway Obstruction
NOSE	Nasal Obstruction Symptom Evaluation scale
NVC	Nasal Valve Collapse
OSA	Obstructive Sleep Apnea
PP	Per Protocol
PRO	Patient-reported Outcomes
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SD	Standard Deviation

VAS	Visual Analogue Scale

2 INTRODUCTION

Nasal airway obstruction (NAO) can be caused by several independent or concomitant factors including septal deviation, enlarged turbinates or a weakened nasal lateral wall, leading to nasal valve collapse (NVC). Surgery to strengthen the lateral wall has been shown to significantly improve the quality of life for subjects suffering from NAO, however current procedures are generally invasive and have the potential to permanently alter the patient's appearance. This study utilizes a minimally invasive technique to address NVC by supporting the nasal lateral wall cartilage with an absorbable implant. Spirox has developed the Latera Absorbable Nasal Implant and Delivery Device, to enable a less invasive alternative to current surgical approaches used to support weak lateral wall cartilage.

The study purpose is to evaluate the Latera Absorbable Nasal Implant (Latera Implant) versus Sham Control in subjects with nasal valve collapse due to or primarily due to insufficient cartilaginous support of the lateral nasal wall.

2.1 Design, Treatments and Visits

This clinical study is to evaluate the performance of the device in comparison to no intervention, in this case, a "sham treatment". The study design and methods have been developed using input from previous and ongoing studies with the Latera Implant. This randomized, single-blind, sham-controlled, multicenter study of the Latera Nasal Absorbable Implant will enroll subjects with nasal valve collapse due to or primarily due to insufficient cartilaginous support of the lateral nasal wall into one of two treatment arms, a) implantation with the Latera device or b) Sham Control.

There are three treatment arms:

Latera Treatment Arm

Subjects in the active treatment arm will receive the Latera Implant, a PLLA-PDLA copolymer, delivered to the region of the lateral nasal wall cartilages to provide support. It is placed using standard, minimally invasive techniques and is absorbed over a period of approximately 18 months.

Sham Control Arm

Subjects in the Sham Control arm will undergo the same preoperative assessments as those in the Latera Treatment arm up to and including anesthesia for the implant and cannula insertion into the nasal lateral wall, however, no implant will be placed.

Randomization to the Sham Control Arm will be assessed by the independent Data Monitoring Committee (DMC) at each interim analysis time point. If the pre-determined efficacy, safety, or operational criteria are met, randomization may be terminated.

Crossover Subjects

Subjects will be unblinded after the 3-month assessment is complete. For subjects in the Sham Control arm, the subject will be treated with the Latera Implant if they still meet all eligibility criteria. The subject's data will continue to be recorded thereafter through 24 months post-implant. If the subject does not meet the eligibility criteria, he/she will exit the study.

For subjects randomized to the Latera Implant arm, the follow-up evaluations will be at 7 days, 30 days, 3 months, 6 months, 12 months, 18 months, and 24 months post-procedure. For subjects randomized to the Sham Control arm, the follow-up evaluation will be at 7 days, 30 days, and 3 months.

2.2 Objectives

The primary objective of the LATERA RCT is to demonstrate the superiority of the Latera Implant to improve nasal breathing, compared with a Sham Control procedure.

2.3 Subject Population

There are up to 150 subjects who will be recruited from sites' existing patient populations that are seeking treatment for nasal airway obstruction including nasal valve collapse. Anatomical considerations may impact the race distribution (e.g. some races may be less prone to lateral wall collapse), but the intention is to enroll all eligible subjects.

Study brochures containing information on study participation and the Latera Implant may be provided to the sites, as well as posters that may be displayed either as hard copies or

electronically on computer monitors in the office. All recruiting materials must be approved by the appropriate Investigational Review Board (IRB) (central or local).

2.3.1 Patient Eligibility

The trial will include up to 150 subjects at up to 15 investigational sites, male and female, aged 18 and above who are appropriate candidates for nasal intervention with the Latera Implant.

Inclusion Criteria

Subjects must meet the following criteria to be included in the study:

1. Adults aged 18 and above;
2. Understands and provides written informed consent;
3. Stated willingness to comply with all study procedures, post-treatment care and availability for the duration of the study follow up of 2 years;
4. In good general health as evidenced by medical history;
5. NOSE score ≥55;
6. Dynamic bi-lateral nasal wall insufficiency as confirmed by Positive Modified Cottle Maneuver;
7. Nasal and facial anatomy appropriate to receive the Latera Implant;
8. Documented failure of benefit after at least 4 weeks of conservative medical management, including, for example, antihistamines or nasal steroids, evidenced by lack of efficacy or tolerability.

Exclusion Criteria

Subjects meeting any one of the following criteria are ineligible for study participation:

1. Unable to tolerate or not a candidate for procedures performed under local anesthesia;
2. Pathology other than lateral wall insufficiency (e.g. septal deviation, turbinate or adenoid hypertrophy, polyps, sinusitis, rhinitis) is the primary contributor to airway obstruction;
3. Requires or is anticipated to require any other concurrent nasal procedures (e.g. Functional Endoscopic Sinus Surgery (FESS), rhinoplasty, sinuplasty, septoplasty, or turbinate reduction) outside of the index procedure within 12 months after the index procedure;

4. FESS, sinuplasty, septoplasty, inferior turbinate reduction, or rhinoplasty within the past 6 months;
5. Any other rhinoplasty procedures are planned or planned usage of external dilators within 24 months after the index procedure;
6. Permanent nasal implant of any type (e.g. autologous, homologous, or synthetic graft) or dilator;
7. Presence of concomitant inflammatory or infectious conditions or unhealed wounds in the treatment area (e.g., vestibulitis, vasculitis, active acne),
8. Currently use of chronic systemic steroids or recreational intra-nasal drugs;
9. Currently has cancerous or pre-cancerous nasal lesions, has had radiation in the treatment area, or is currently receiving chemotherapy;
10. History of a significant healing disorders including hypertrophic scarring, or keloid formation;
11. Poorly controlled diabetes mellitus;
12. Known or suspected allergy to PLA or other absorbable implant materials in the Latera Implant;
13. Severe obstructive sleep apnea (OSA) and cannot or is unwilling to refrain from continuous positive airway pressure (CPAP) for 4 weeks post-procedure, in agreement with the treating physician;
14. Female subjects, of child bearing potential, known or suspected to be pregnant or are lactating;
15. Any other presenting condition that, in the medical opinion of the investigator, would disqualify the subject from the study.

2.4 Sample Size Considerations

The primary analysis will be based on a 1-sided binomial test of proportions. Assuming a Sham Control response rate of 40% and a Latera Treatment response rate of 70% at month 3, a maximum sample size of 124 evaluable subjects is required for 90% power and preserving a 2.5% (one-sided) type I error rate. This study is designed as a group-sequential trial with 2 interim analyses and a final analysis. The stopping boundaries are derived using the Triangular

Method[1], which is a special case of Unified Family (Power) boundaries. A shape parameter of 0.65 will be used, and both boundaries will be non-binding.

Interim analyses will occur when approximately 20 and 60 total subjects have been enrolled and have had their 3-month assessment. The Latera response rate has been observed at approximately 85% in single-arm trials and is likely to be slightly lower in a single blind setting. Thus, 70% is likely a conservative estimate. The placebo response rate observed in pain trials is approximately 30%, and may be even higher in a procedural setting, so 40% was selected for the control response rate used in the study design calculations.

Assuming up to a 10% drop out rate and a 10% potential retrieval rate, the sample size is adjusted to enroll up to 150 subjects to assure adequate sample size with 3 months follow up.

2.5 Randomization

This is a randomized trial and subjects are randomized into Sham Control Arm or Latera Implant Treatment Arm, after anesthetic is delivered. Treatment assignment will remain blinded to the study subject through the 3-month assessment.

The graphic below illustrates the flow of subjects in the study and potential points of study exit relative to the analysis plan (**Figure 1**, below). Subjects are enrolled at the time of consent, and the ITT population is comprised of all consented subjects. Subjects are randomized after the anesthetic is administered. Subjects in the active treatment group receive the Latera implant, while those in the Sham Control group will undergo insertion of the cannula into the nasal lateral wall, but the Latera will not be implanted. The mITT population includes all subjects who were randomized, irrespective of whether they received the study device. The PP population includes those subjects who received the study device or the Sham Control and who are followed through 3 months.

[1] Kittelson JM, Emerson SS. A unifying family of group sequential test designs. Biometrics. 1999 Sep;55(3):874-82.

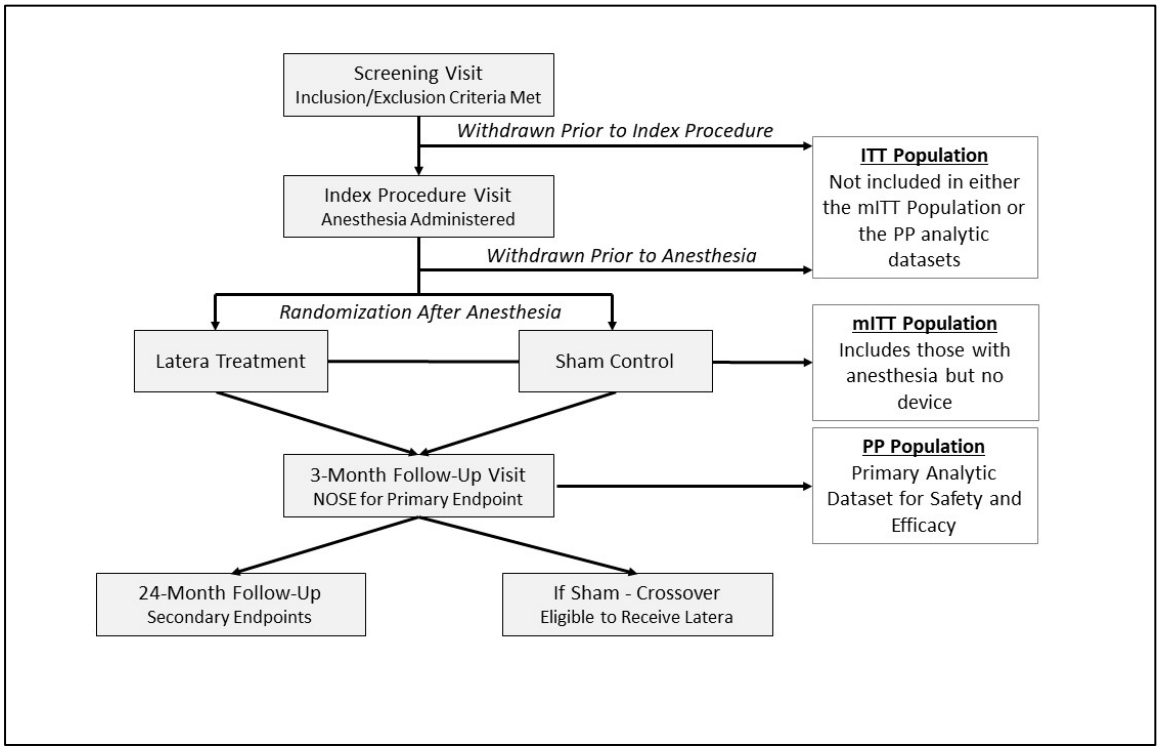

Figure 1. Flow of subjects in the trial and definitions of the ITT, mITT and PP populations.

2.6 Study Activities and Assessments

Table 1. Schedule of Study Activities

Visits (± Windows)	Screening/Baseline	Index Procedure - Day 0	Visit 3 - 7 ± 2 days	Visit 4 - 30 ± 7 days*	Visit 5 - 90 ± 14 days*	Cross-Over Screening (for Sham Control Subjects	Cross-Over Procedure (for Sham Control Subjects	Visit 6 – 6 month ± 14 days	Visit 7 – 12 months ± 14 days	Visit 8 – 18 months ± 14 days	Visit 9 – 24 months ± 14 days	
Informed Consent	X					X						
Nasal Medical History & Exam	X											
Eligibility	X											
Lateral Wall Motion Video	X				X			X				
Turbinate Hypertrophy and Septum Assessment	X											
Nasal Obstruction Evaluation (NOSE)	X			X	X	X		X	X	X	X	
Sleep Assessments (ESS & PSQI)	X			X	X	X		X	X	X	X	
Photograph – Planning Images	X											
Pregnancy Test	X											
Nasal Breathing Assessment (VAS)	X		X	X	X	X		X	X	X	X	
Randomization		X										
Procedure Logistics		X					X					
Relevant Medications	X		X	X	X			X	X	X	X	
Adverse Events	X	X	X	X	X		X	X	X	X	X	

Note: *For Sham Control Subjects that cross-over after 90 day assessment for primary endpoint and meet study eligibility for a Latera Implant, per the protocol, will have repeat 30 and 90 day visits, in addition to all the other visits.

The schedule of study activities (Table 1, above) specifies the assessments that will be completed at each follow up visit. Below are the descriptions of the Patient Reported Outcomes (PRO)_ assessments and Lateral Wall Motion Video assessment:

Nasal Obstruction Symptom Evaluation (NOSE) scale

The Nasal Obstruction Symptom Evaluation (NOSE) scale is a PRO instrument that will be administered to capture subject perception of the degree of nasal airway patency.

The NOSE scale is a validated instrument, developed by the American Academy of Otolaryngology—Head and Neck Surgery (AAO-HNS), and has been used in several clinical trials. The scale is brief, easy to complete, and is an important tool for pre- and post-intervention evaluation of symptoms in subjects with nasal obstruction.

Subjects will be asked: "Since your last follow up visit, how much of a problem were the following conditions for you?" Specifically, subjects will be asked to rate their perceptions on the Likert scale with respect to the following characteristics:

- Nasal congestion or stuffiness
- Nasal blockage or obstruction
- Trouble breathing through my nose
- Trouble sleeping
- Unable to get enough air through my nose during exercise or exertion

Subjects will rate their responses using a Likert scale with response options 0, 1, 2, 3 or 4, as follows:

- Not a Problem
- Very Mild Problem
- Moderate problem
- Fairly Bad Problem
- Severe problem

The responses are rated along the continuum, with a rating of "0 – not a problem" indicating no problem breathing, with a completely free flow of air through the nasal airway; "1 – very mild problem", with only slight obstruction in airflow; "2 – moderate problem", with mouth breathing

considered easier; "3 – fairly bad problem," with considerable obstruction to airflow; and a rating of "4 – severe problem," with complete blockage and obstruction of the nasal passageway, where the subject cannot breathe through the nose and can only mouth breathe.

Nasal Breathing Assessment

A subject's perception of breathing cannot be quantitatively measured but exists on a continuum from the subject perspective. The VAS is a PRO instrument that will be used to capture subjects' perception of their ability to breathe through the nose, allowing subjects to indicate the degree of breathing difficulty (or ease) they are currently experiencing.

Operationally, the VAS is a horizontal line, 100mm in length, anchored by word descriptors at each end, as illustrated in **Figure 2**. The subject will mark on the line the point that they feel represents their perception of their current state and the electronic PRO form will automatically calculate the score for VAS. The VAS score is determined by measuring in millimeters from the left end of the line to the point marked by the subject.

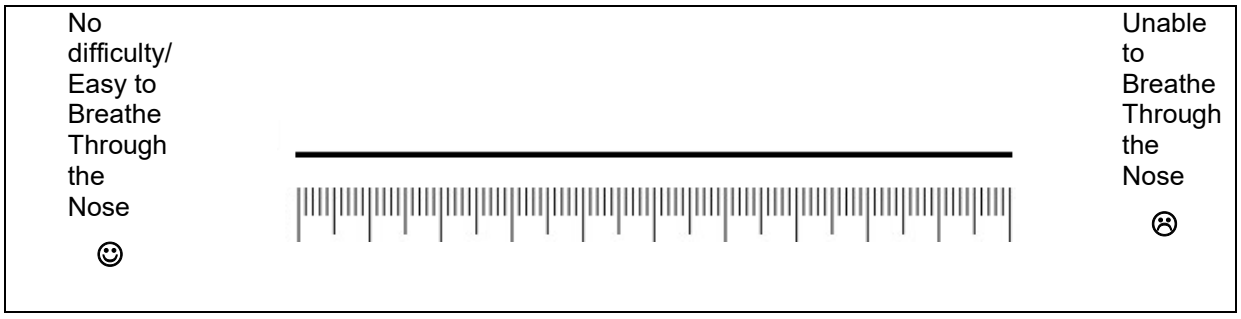

Figure 2: Visual Analog Scale. Left side represents 0 mm and right side represents 100 mm.

Epworth Sleepiness Scale (ESS)

The Epworth Sleepiness Scale (ESS) is a widely-used PRO in the field of sleep medicine as a subjective measure of a patient's sleepiness.

The test is a list of eight situations that evaluate a subject's tendency to become sleepy on a scale of 0 (no chance of dozing), to 3 (high chance of dozing). The scale estimates whether the subject is experiencing excessive sleepiness that may possibly require medical attention.

Subjects will be asked: "How sleepy are you? How likely are you to doze off or fall asleep in the following situations?" They will be asked to rate their chances of dozing off, not just feeling tired and even if they have not done some of these things recently to try and determine how they would have affected them. For each situation, subjects will be asked to decide if they would have:

- No chance of dozing 0
- Slight chance of dozing 1
- Moderate change of dozing 2
- High chance of dozing 3

Using the above scale, subjects will write down the corresponding choice to the following situations:

- Sitting and reading
- Watching TV
- Sitting inactive in a public place (e.g. a theater or a meeting)
- As a passenger in a car for an hour without a break
- Lying down to rest in the afternoon when circumstances permit
- Sitting and talking to someone
- Sitting quietly after lunch without alcohol
- In a car, while stopped for a few minutes in traffic

The responses to the above situations is totaled for a final composite score.

Pittsburg Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) is a questionnaire used to assess sleep quality over the previous 1-month. The measure consists of 19 individual items, creating 7 components that produce one total score. Component scores consist of subjective sleep quality, sleep latency (i.e., how long it takes to fall asleep), sleep duration, habitual sleep efficiency (i.e., the percentage of

time in bed that one is asleep), sleep disturbances, use of sleeping medication, and daytime dysfunction.

Lateral Wall Motion Video Assessment

Video results from Endoscopy assessments will be captured and transferred to Spirox. The videos will be de-identified prior to transfer to an independent reviewer. The independent reviewer will make assessment of endoscopic lateral wall insufficiency score (1, 2 or 3).

2.7 Retrievals and Re-Implantation

There is a low incidence of where the Latera Implant is removed, post-implant nasal manipulation by a patient. If this occurs, the Latera Implant may be re-implanted, within 30 days, of the removal at the discretion of the investigator, continuing the follow up schedule from the Index Procedure. An addendum to the Informed Consent is required to be reviewed and signed by subject. See Section 4.1.1 for analysis implications.

3 ENDPOINT DEFINITIONS

3.1 Primary Endpoint

The primary endpoint of the study is the Responder Rate, assessed 3 months post-procedure in the per protocol population of the Latera and Sham arms. Responder Rate is defined as the proportion of subjects with at least one (1) NOSE class improvement or at least 20% NOSE score reduction.

3.2 Secondary Endpoints

The following endpoints comprise the secondary endpoints and the time points that each will be assessed:

- Responder Rate at 7 days, 30 days, and 6, 12, 18 and 24 months;

- Frequency of procedure-related adverse events (non-serious and serious), assessed at the index procedure and through the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points;
- Frequency of device-related adverse events (non-serious and serious), assessed at the index procedure and through the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month and 24-month follow-up time points;
- NOSE score, change from baseline, to the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points;
- Visual Analogue Scale (VAS), change from baseline, to the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points;
- Endoscopic lateral wall insufficiency (LWI) score as assessed by independent reviewers, change from baseline to the 3-month and 6-month time points;
- Epworth Sleepiness Scale (ESS) score and Pittsburg Sleepiness Quality Index, change from baseline, to the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points.

All secondary endpoints will be assessed for the Sham Control Arm at two additional visits. The Sham arm will have follow-up visits at 7 days, 30 days, and 3 months after the initial treatment/randomization, and 30 days, 3 months, 6, months, 12 months, 18 months, and 24 months in subjects that cross over.

4 ANALYSIS POPULATIONS

4.1.1 Study Population

Adult male and female patients presenting with symptomatic nasal valve collapse due to or primarily due to insufficient cartilaginous support of the lateral nasal wall will be enrolled, if they meet all eligibility criteria. The study will include up to 150 subjects enrolled at up to 15 investigational sites in the US. The maximum enrollment per site will be 20% of the total enrollment number therefore a 30-subject maximum per site. Initial site enrollment will be limited to 20-subjects and written permission from Sponsor will be required to enroll up to 30 subjects maximum.

The following populations will be analyzed in the study:

Intention-to-Treat (ITT) Population: The ITT analysis will be performed on all enrolled subjects in the study, irrespective of adherence with the entry criteria, treatment actually received, subsequent withdrawal, or deviation from the Clinical Investigational Plan.[2] Subjects who are enrolled but not randomized will not be included in the analytic datasets.

Modified Intention-to-Treat (mITT) Population: The mITT Population will comprise of ITT subjects who received anesthetic, regardless of whether they received the Latera Implant or underwent the Sham procedure. Subjects who are in the mITT population assigned to the Latera Implant Arm but who do not receive the device are followed for 3 months for safety-related endpoints and will be exited after 3-month follow-up assessments.

Per Protocol Population: The PP analysis will be performed on randomized subjects a) in the Sham Control Arm, and b) in the Latera Implant Arm who were treated and who had NOSE score assessments within the 3-month follow-up clinic visit window. The primary endpoint will be evaluated on this population. Subjects will be analyzed according to the procedure actually received in the event of mis-randomization. Unblinded subjects in either arm prior to the 3-month time point are not included in the PP analysis.

The point of enrollment into the study occurs when the subject signs the Informed Consent Form (ICF) and meets all eligibility criteria.

4.2 Unblinding and Crossover Subjects

All subjects will be un-blinded after the assessment is complete at the 3-month time point. If the subject is in the Sham Control group, a Latera implant may be implanted if they meet the eligibility criteria. Subjects who are un-blinded (within either treatment arm) prior to the 3-month time point will be excluded from the per protocol analysis dataset for primary and secondary analysis endpoints. However, data will continue to be recorded through the 24-month follow-up time point for these subjects.

[2] Fisher LD, Dixon DO, Herson J, Frankowski RK, Hearon MS, Pearce KE. Intention to treat in clincial trials. In: Pearce KE, ed. Statistical issues in drug research and development. New York: Marcel Dekker; 1990:331-50

Sham Control subjects may not receive a nasal implant prior to the 3-month primary endpoint assessment.

Cross-over subjects will have 2 additional visits as they will begin follow-up time points again once they have received the Latera Implant.

5 STATISTICAL ANALYSES

5.1 General Considerations

All statistical programming will be performed using SAS® version 9.4 or above.

Descriptive summary statistics will be presented for all data points. When data are not available, or partially available data summaries will be based on available data. For continuous variables, number of observations, mean, standard deviation (SD), median, minimum, maximum and, for some variables, the 95% confidence interval (CI) will be presented. For categorical variables, numerator, denominator, percentage (%) and for some variables, the 95% CI will be presented. In both cases the difference between treatment and control as well as 95% CI of the difference may be presented. When presented, the 95% CI will be calculated using the exact method.

5.2 Primary Endpoint Analysis

The primary endpoint of the study is defined as the proportion of responders at 3 months following the index procedure. Responder is defined as a subject that has at least one (1) NOSE class improvement or at least 20% NOSE score reduction from baseline to 3 months post treatment.

The treatment (p_T) and control (p_C) will be compared using a 1-sided test for superiority of the following hypothesis:

$$H_0: \ p_T \leq p_C \text{ versus } H_A: \ p_T > p_C$$

Where p is the proportion of subjects with a positive response on the primary endpoint assessment. Successful rejection of the null hypothesis will be taken as evidence the device provides reasonable assurance of effectiveness.

A 1-sided binomial test of proportions will be used to compare treatment groups.

5.3 Secondary Endpoint Analyses

The following endpoints comprise the secondary endpoints and the time points that each will be assessed:

- Responder Rate at 7 days, 30 days, and 6, 12, 18 and 24 months;
- Frequency of procedure-related adverse events (non-serious and serious), assessed at the index procedure and through the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points;
- Frequency of device-related adverse events (non-serious and serious), assessed at the index procedure and through the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month and 24-month follow-up time points;
- NOSE score, change from baseline, to the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points;
- Visual Analogue Scale (VAS), change from baseline, to the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points;
- Endoscopic lateral wall insufficiency (LWI) score as assessed by independent reviewers, change from baseline to the 3-month and 6-month time points;
- Epworth Sleepiness Scale (ESS) score and Pittsburg Sleepiness Quality Index, change from baseline, to the 7-day, 30-day, 3-month, 6-month, 12-month, 18-month, and 24-month follow-up time points.

All secondary endpoints will be assessed for the Sham Control Arm at two additional visits. The Sham arm will have follow-up visits at 7 days, 30 days, and 3 months after the initial

treatment/randomization, and 30 days, 3 months, 6, months, 12 months, 18 months, and 24 months in subjects that cross over.

A 2-sided binomial test of proportions will be used to compare treatment groups for binary endpoints. A two-sample t-test will be used for continuous endpoints.

5.4 Additional Analyses

5.4.1 Demographics and Baseline Characteristics

Demographic, medical history and other clinically relevant baseline variables, as well as safety data, excluding AEs and SAEs, will be summarized by treatment using descriptive statistics (i.e. number of observations available, mean, standard deviation, minimum, and maximum for continuous variables and counts and percentages for qualitative variables). Dichotomous variables will be evaluated using Fisher's exact tests. Categorical variables will be evaluated by Cochran-Mantel-Haenszel (CMH) Modified Ridit Scores, i.e. CMH of general association for nominal variables and CMH of row mean score for ordinal variables. Continuous variables will be evaluated by a two-sample t-test.

5.4.2 Subgroup Analyses

Additional subgroup analyses are planned for gender. These additional analyses will be conducted as feasible and may be used in combination with all other analyses for drawing conclusions regarding the safety and effectiveness of the investigational device.

5.4.3 Changes to Planned Analyses

Any changes to the planned analyses will be documented as amendments to the Statistical Analysis Plan (SAP) and in the study report.

5.4.4 Interim Analyses

There will be two planned interim analyses and one final analysis for this study (**Table 1**, below). The stopping boundaries are derived using the Triangular Method [see footnote 1], which is a special case of Unified Family (Power) boundaries. A shape parameter (τ) of 0.65 will be used, and both boundaries will be non-binding. Interim analyses will occur when approximately 22 and 60 total subjects have been enrolled with 3-month assessments. The stopping boundaries are

moderately aggressive for both effectiveness and futility to reduce the number of subjects enrolled if clear evidence of effectiveness or futility becomes apparent during the trial. When the design assumptions are correct (40% versus 70%), the chance that the trial will stop for effectiveness at the second interim analysis is about 1 in 3. However, if the Latera rate is 85% as seen in the single arm studies, the chance of stopping for effectiveness at the second interim analysis is better than 8 in 10. While all evaluable subjects enrolled at the time of the interim analysis would be followed through 3 months and included in the primary endpoint, an early termination at the second interim analysis could save 20-30% of the total sample size, depending on the enrollment rate.

Interim analyses will be carried out by an independent statistician and reviewed by the independent DMC. The interim analyses will use an alternative to the critical ratio test given below and described by Fleiss[3]. The test results will be compared to the triangular group sequential boundaries. Additional operational details will be found in the DMC Charter.

$$z' = \frac{|p_2 - p_1| - \frac{1}{2}(1/n_{1.} + 1/n_{2.})}{\sqrt{\frac{p_1 q_1}{n_{1.}} + \frac{p_2 q_2}{n_{2.}}}}.$$

Table 1. Group Sequential Trial Design – Boundary Information

Analysis	Evaluable (n)	Boundaries (p-value)		Cumulative Error Spending	
		Futility	Efficacy	Type II	Type I
Interim 1	22	0.7921	0.0005	0.0106	0.0005
Interim 2	60	0.2139	0.0090	0.0548	0.0093
Final	124	0.0194	0.0194	0.1000	0.0250

[3] Fleiss JL, B Levin, MC Paik. Statistical Methods for Rates and Proportions, 3rd Ed. John Wiley & Sons, 2003, (Ch. 3, p 61).

SAS Code for Nonbinding Beta Boundary

```
proc seqdesign altref=0.3 errspend          BOUNDARYSCALE=pvalue;
onesidederrorspending: design nstages=3 info=cum(3.6 8 20)/*3/20,
10/20=1/2, 1*/
method(alpha)=tri(tau=.65)
method(beta)=tri(tau=.65)
alt=upper stop=both (betaboundary=nonbinding)
alpha=0.025 beta=0.10;
samplesize model = twosamplefreq (nullprop=0.4 test=prop);
ods output Boundary=Bnd_Prop_pt5 ;
run;
```

5.5 Missing or Incomplete Data

Every effort will be made to minimize the amount of missing data. Recognizing the difficulty of avoiding some missing data, however, data imputation methods with sensitivity imputation analyses will be conducted as secondary analyses for the primary endpoint in the ITT and mITT populations. The robustness of the multiple imputation-based sensitivity analyses for the primary outcome will be tested with a tipping point analysis encompassing all possible imputation outcomes[4].

As a secondary sensitivity analysis in the ITT and mITT populations, multiple imputation of the primary endpoint will be carried out using the logistic regression approach for a dichotomous outcome with PROC MI in SAS for patients missing data for the 3-month assessment. The following baseline variables will be included in the imputation model as covariates:

- Age
- Gender
- Body Mass Index (BMI)
- Nose Length
- Nose Height
- Nose Width
- Inferior Turbinate Hypertrophy
- Septal Deviation

[4] Yan X, Lee S, Li N. Missing data handling methods in medical device clinical trials. J Biopharm Stat, 19, 1085-98, 2009.

- NOSE Score
- Breathe VAS Score
- Sleep Apnea
- Nasal Surgery
- Epworth Sleepiness Scale

Five data sets will be imputed from these covariates and will mimic different realizations of the missing data. Within each imputed (complete) dataset for the primary endpoint, the proportion experiencing the endpoint will be statistically compared between treatment group using the two-sample z-test. For the primary endpoint, the numerator of the test statistic (the numerator is the point estimate of the treatment difference for the primary endpoint) and its standard error (the pooled standard error of treatment difference for the primary endpoint) will be pooled across the 5 data sets using established variance-adjustment methods (e.g., via PROC MIANALYZE in SAS) to create one overall numerator and denominator. From these, an overall test statistic for the endpoint and its associated p-value will be calculated for the imputed data.