Statistical Analysis Plan

InspIRE (Wave I)

Pulsed Field Ablation (PFA) System for the Treatment of Paroxysmal Atrial Fibrillation (PAF) by Irreversible Electroporation (IRE)

Protocol # BWI_2019_08 (v 1.0)

Version 2.0

April 6, 2022

Sponsor: BIOSENSE WEBSTER, INC.
31 Technology Dr.
Irvine, CA 92618
USA

History of Changes:

Version— Date	**Description**
V 1.0 - 13 May 2021	Original document
V 2.0 – 06 April 2022	1. The definition of the Per Protocol analysis set was modified to match the definition in the inspIRE Wave II SAP. 2. Clarification was provided on the derivation of acute procedural failure for subjects with non-study catheter use and failure of PFA delivery. 3. Subjects without failure with repeat ablation of PVs with a non-study catheter during the blanking period were specified to be excluded from long-term effectiveness analyses. 4. The analysis sets specified for analyses of endpoints and any related sensitivity analyses were modified to match at minimum what is specified in the inspIRE Wave II SAP. 5. Clarification was provided on the handling of missing data for the primary effectiveness endpoint.

Contents

List of Abbreviations

AAD	Antiarrhythmic Drug
AF	Atrial Fibrillation
AFEQT	Atrial Fibrillation Effect on QualiTy-of-life
AFL	Atrial Flutter
AT	Atrial Tachycardia
CEC	Clinical Events Committee
CT	Computed Tomography
DMC	Data Monitoring Committee
FAS	Full Analysis Set
HRQoL	Health Related Quality of Life
IRE	Irreversible Electroporation
mITT	Modified Intent-To-Treat
MMSE	Mini Mental State Examination
MRA	Magnetic Resonance Angiogram
MRI	Magnetic Resonance Imaging
mRS	Modified Rankin Scale
NIHSS	National Institutes of Health Stroke Scale
NSC	Non-study catheter
PAE	Primary Adverse Event
PAF	Paroxysmal Atrial Fibrillation
PFA	Pulsed Field ablation
PP	Per-Protocol
PV	Pulmonary Vein
PVI	Pulmonary Vein Isolation
QOL	Quality of Life
RF	Radiofrequency

SADE	Serious Adverse Device Effects
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan

1 STUDY DESIGN

This clinical investigation is an interventional, prospective, single arm, multi-center, pre-market clinical evaluation of the IRE ablation system (Circular IRE Catheter and IRE Generator) to demonstrate acute safety and long-term effectiveness when compared to predetermined performance goals. The study will enroll subjects with drug refractory, symptomatic paroxysmal atrial fibrillation (PAF) who are candidates for catheter ablation in two sequential waves, including 1) Wave I subjects who will undergo the index procedure and additional neurological, pulmonary vein (PV), and esophageal assessments, and 2) Wave II roll-in and main phase subjects who will undergo the index procedure and same follow-up schedule as Wave I subjects, excluding the additional neurological, post-procedural PV (unless symptomatic), and esophageal assessments. This statistical analysis plan (SAP) describes the analyses that will be performed for subjects in Wave I. A separate SAP will describe the analyses for roll-in and main phase subjects in Wave II.

For the purpose of characterization of safety and to provide preliminary estimates for safety and acute effectiveness of the IRE system, Wave I of the study will enroll up to 40 subjects in a maximum of 7 sites. The subjects of Wave I will meet all eligibility criteria, including the additional exclusion criteria specific to Wave I, and will be evaluated at 7 days, 1, 3, 6, and 12 months following the index procedure. Wave II enrollment will be initiated when all Wave I subjects have reached 7 days follow up unless the Sponsor deems otherwise per sections 25 (Study Suspension or Termination) and 6.4.2 (Minimization of Risk) of the study protocol. A Clinical Events Committee (CEC) will be implemented to adjudicate primary safety endpoint events for Wave I and Wave II. The CEC will operate as described in the CEC Charter.

2 TREATMENT ASSIGNMENT

All subjects will be treated with the IRE ablation system, whose components consist of

- Multi-channel IRE generator (D-1417-01-I)
- Multi-electrode circular IRE catheter (D-1412-01-SI)
- Related components and accessories

3 RANDOMIZATION AND BLINDING PROCEDURES

This is a non-randomized trial with all subjects receiving treatment with the IRE ablation system. Therefore, masking of treatment assignment for operators and subjects will not be performed. However, in order to minimize operational bias, screening logs will be maintained at sites to confirm consecutive eligible subjects are considered for participation in the study.

4 LEVEL OF SIGNIFICANCE

No formal statistical hypothesis or inferential statistics will be formulated or performed in Wave I of the study.

5 ANALYSIS SETS

For the analysis of study endpoints, the analysis sets defined in the following will be used:

- **Full Analysis Set (FAS):** The FAS analysis set will consist of all enrolled Wave I subjects who have had insertion of the study catheter (with or without delivery of pulse field ablation [PFA]).

- **Modified Intent-To-Treat (mITT) Analysis Set**: The mITT analysis set will consist of enrolled Wave I subjects who have had insertion of the study catheter and meet eligibility criteria.

- **Per Protocol Analysis Set (PP)**: The PP analysis set will consist of Wave I subjects who satisfy the following criteria:
    - Have undergone ablation with the IRE ablation system.
    - Are treated for the study-related arrhythmia.
    - Are without major protocol deviations that would affect the scientific integrity of the primary safety and effectiveness data, including but not limited to the following:
        - Subjects found not meeting eligibility criteria
        - Use of the IRE ablation modality outside of the PV region
        - Failure to check for entrance block for each targeted PV after adenosine/isoproterenol challenge

      Eligible subjects treated with commercially available catheters for non-PV triggers that may arise during the index procedure or during repeat procedures occurring during the blanking period will be followed for the full duration of follow-up and will be evaluable for all endpoints in order to preclude potential bias that would arise from exclusion of these subject profiles from endpoint estimates.

6 SAMPLE SIZE JUSTIFICATION

As Wave I of the study is a safety characterization phase, a sample size of 40 subjects is intended to delineate safety and provide preliminary estimates for safety and acute effectiveness of the

IRE system. The maximum enrollment of 40 subjects will provide greater than 90% probability of observing at least one primary adverse event (PAE), assuming the PAE rate is 7% and that 35 of the enrolled subjects have the study catheter inserted.

7 STATISTICAL ANALYSIS METHODS

Wave I subject data will be analyzed separately (not part of the primary hypothesis testing). No formal statistical inference will be made and all analyses will be descriptive. The following confidence intervals may be constructed and presented for clinical relevance only:

- Exact two-sided 95% confidence intervals for safety and effectiveness endpoints
- Two-sided 95% confidence intervals using Greenwood's variance for Kaplan-Meier estimates for effectiveness endpoints

For all effectiveness endpoints, subjects who are discontinued (no energy [PFA] delivered with the study catheter) due to:

- IRE system related reasons will be considered acute effectiveness failures
- Non IRE system related reasons (e.g. other equipment or anatomy that precludes treatment with the study catheter or commercially available catheter) will be excluded from the effectiveness analysis.

7.1 General Conventions

Standard descriptive summaries for continuous data include the number of observations with data, mean, standard deviation, median, 25th percentile, 75th percentile, minimum, and maximum values. For categorical data, the count and percent will be provided. Percentages will be based on the number of subjects without missing data.

7.2 Subject Disposition

Disposition and accountability of the study subjects will be summarized descriptively for the subject categories defined in section 10.3 of the study protocol.

7.3 Demographic and Baseline Characteristics

Subject demographics, medical history, previously failed antiarrhythmic drugs (AADs), active AAD use at baseline, and other baseline data will be summarized descriptively for subjects in the PP, mITT, and FAS analysis sets.

7.4 Analysis of Primary Endpoints

7.4.1 Acute Safety

The primary safety endpoint for Wave I is the occurrence of any PAE within 7 days of the initial mapping and ablation procedure. The definition of PAE can be found in section 8.2.1 of the study protocol.

The PAE rate within 7 days of the initial ablation will be summarized descriptively in the mITT analysis set. The number of subjects with primary safety events, the total number of primary safety events, and the percentage of subjects with primary safety events will be presented.

The primary safety endpoint will also be analyzed in the FAS analysis set as a sensitivity analysis.

7.4.2 Long-Term Effectiveness

The primary effectiveness endpoint is defined as freedom from documented (symptomatic and asymptomatic) atrial arrhythmia (atrial fibrillation [AF], atrial tachycardia [AT] or atrial flutter [AFL]) recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of a non-study catheter for PV isolation, or failure to have PFA delivery with the study catheter due to IRE system malfunctions) will also be considered a long-term effectiveness failure. Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis.

The primary effectiveness endpoint will be summarized descriptively in the PP analysis set. The number and percentage of subjects free from primary effectiveness failure during the evaluation period will be presented.

The following additional analyses will be performed at the time of the effectiveness analysis:

- **Sensitivity to Analysis Set**

  The primary effectiveness endpoint will be analyzed in the mITT and FAS analysis sets as sensitivity analyses.

- **Kaplan-Meier Analysis**

  Kaplan-Meier estimates and plots will be used to characterize the time to the first primary effectiveness event, including acute procedural failure and documented symptomatic/asymptomatic AF, AT, or AFL recurrence following the initial ablation

procedure, in the FAS, mITT, and PP analysis sets. The probability of freedom from primary effectiveness failure at each follow-up timepoint post blanking will be presented.

7.4.3 Handling of Missing Data

Missing data will be queried for reasons and will not be imputed.

For the primary safety analysis, if a subject's follow-up time is less than 3 months and the subject has not had a PAE, then that subject will be excluded from the primary safety analysis.

For the primary effectiveness analysis, subjects without an effectiveness failure and who (1) do not have full 12 months of follow-up and/or sufficient follow-up duration for the primary effectiveness endpoint (i.e., with at least 335 days of follow-up or arrythmia monitoring post the index procedure), or (2) are missing all remote arrythmia and holter monitoring records as required by protocol will be considered missing primary effectiveness endpoint and will be excluded from the primary effectiveness analysis.

7.4.4 Site Heterogeneity

Site heterogeneity will not be tested for Wave I.

7.5 Analysis of Secondary Endpoints

Descriptive statistics will be provided on all secondary endpoints in the PP, mITT, and FAS analysis sets.

7.5.1 Acute Procedural Success

Acute procedural success is defined as confirmation of entrance block in all clinically relevant targeted PVs after adenosine/isoproterenol challenge. Use of a non-study catheter to achieve PVI and failure to have PFA delivery with the study catheter due to IRE system malfunctions are considered acute procedural failures.

The rate of acute procedural success will be summarized descriptively. The number and percentage of subjects with acute procedural success will be presented.

7.5.2 Freedom from Documented Symptomatic Recurrence

The symptomatic recurrence endpoint is defined as freedom from documented symptomatic atrial arrhythmia (AF, AT or AFL) recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of non-study catheter for PVI, and failure to have PFA delivery with the study catheter due to IRE system malfunctions) will also be considered a failure. Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis.

Kaplan-Meier estimates and plots will be used to characterize the time to first symptomatic AF/AT or AFL recurrence following the initial ablation procedure. The probability of freedom from symptomatic AF/AT or AFL recurrence at each follow-up timepoint post blanking will be presented.

7.5.3 Quality of Life (QOL)

The Atrial Fibrillation Effect on Quality-of-Life (AFEQT™) includes 20 questions on a 7-point Likert scale. Questions 1-18 evaluate Health Related Quality of Life (HRQoL) and questions 19-20 relate to patients' satisfaction with treatment[1]. The first 18 questions are used to calculate the Overall AFEQT score and the subscale scores across three domains as follows:

- AF related Symptoms: Four questions (1 – 4)
- Daily Activities: Eight questions (5 – 12)
- Treatment Concerns: Six questions (13 – 18)

The formula to calculate these scores is as follows:

$$100 - [\frac{(sum\ of\ severity\ for\ all\ questions\ answered - number\ of\ questions\ answered)\ \times 100}{total\ number\ questions\ answered \times 6}]$$

Overall and subscale scores range from 0 to 100. A score of 0 corresponds to complete disability, while a score of 100 corresponds to no disability.

Baseline AFEQT scores and changes from baseline at each timepoint the questionnaire is administered will be summarized descriptively for the following five scores. The overall AFEQT score and subscale scores across study visits will also be plotted.

- Overall AFEQT Score (18 questions)
- Symptom Subscale Score (4 questions)
- Daily Activities Subscale Score (8 questions)
- Treatment Concern Subscale Score (6 questions)
- Treatment Satisfaction Score (2 questions)

The AFEQT questionnaire will only be used in countries with validated languages. These analyses will be conducted for subjects in countries where the AFEQT questionnaires are applied.

7.6 Analysis of Additional Endpoints

Descriptive statistics will be provided on all additional endpoints in the analysis sets specified below.

7.6.1 Procedural data

Procedural data such as total procedure time, mapping time, PFA application time, number of PFA applications by PV and by subject, total fluoroscopy time, study catheter dwell time, ablation settings used, and use of paralytics and anesthesia will be summarized descriptively. These analyses will be conducted separately in the FAS and PP analysis sets.

7.6.2 Additional Safety Endpoints

The following analyses for the additional safety endpoints will be conducted in the mITT and FAS analysis sets and summarized descriptively as the total number of events, number of subjects with events, and percentage of subjects with events:

- Occurrence of individual PAEs from the primary composite
- Occurrence of Serious Adverse Device Effects (SADEs)
- Occurrence of Serious Adverse Events (SAEs) within 7 days (early-onset), 8-30 days (peri-procedural), and >30 days (late onset) of initial ablation procedure, separately for each timeframe
- Occurrence of non-SAEs
- Occurrence of clinically symptomatic severe PV stenosis as documented by computed tomography (CT)/ magnetic resonance angiogram (MRA)

7.6.3 Additional Effectiveness Endpoints

The following analyses for the additional effectiveness endpoints will be conducted separately in the FAS, mITT, and PP analysis sets and summarized descriptively:

- **Ablation by non-study catheter for PV isolation**: The number and percentage of subjects and number and percent of clinically relevant targeted PVs ablated by a non-study catheter (NSC) for PVI.
    - Rate of ablation by NSC for PVI among all clinically relevant targeted pulmonary veins:

$$= \frac{Number\ of\ PVs\ ablated\ by\ NSC}{Number\ of\ ablated\ PVs}$$

    - Rate of ablation by NSC for PVI among subjects:

$$= \frac{Number\ of\ subjects\ \ with\ at\ least\ one\ PV\ ablated\ by\ NSC}{Total\ number\ of\ subjects\ ablated\ for\ PVI}$$

- **Acute PV reconnection:** The number and percentage of subjects and number and percentage of targeted PVs where acute reconnection was identified by adenosine/isoproterenol challenge.
    - Rate of acute PV reconnection among targeted veins:

$$= \frac{Number\ of\ PVs\ with\ reconnection\ after\ adenosine/isoproterenol\ challenge}{Total\ number\ of\ PVs\ undergoing\ adenosine/isoproterenol\ challenge}$$

    - Rate of acute PV reconnection among subjects:

$$= \frac{Number\ of\ subjects\ with\ PV\ reconnection\ in\ at\ least\ one\ PV\ after\ adenosine/isoproterenol\ challenge}{Total\ number\ of\ subjects\ undergoing\ adenosine/isoproterenol\ challenge\ in\ all\ PVs}$$

- **Repeat ablation during the 12-month follow-up period**
    - **Occurrence of repeat procedures:** Kaplan-Meier estimates and plots will be used to characterize the time to the first repeat ablation procedure for study related arrythmia. The probability of freedom from repeat ablation for study related arrythmia at each follow-up timepoint post index procedure will be presented. Additionally, the total number of repeat procedures, number of subjects undergoing a repeat procedure, and percentage of subjects undergoing a repeat procedure will be presented by type of arrhythmia treated (overall (all arrhythmias), AF and AFL or AT (left atrium)), and by timing of the repeat procedure (during the blanking period (Day 1-90) and during the evaluation period (Day 91-365)).
    - **PV reconnection:** The number and percentage of subjects and number and percentage of targeted PVs at the repeat procedure where PV reconnection was observed
        - Rate of PV reconnection among previously isolated veins:

$$= \frac{Number\ of\ previously\ isolated\ PVs\ with\ reconnection\ identified\ at\ repeat}{Total\ number\ of\ previously\ isolated\ veins\ in\ subjects\ undergoing\ repeat\ ablation}$$

- Rate of PV reconnection among subjects:

$$= \frac{Number\ of\ subjects\ with\ PV\ reconnection\ identified\ at\ repeat\ procedure\ in\ at\ least\ 1\ previously\ isolated\ vein}{Total\ number\ of\ subjects\ undergoing\ repeat\ ablation\ procedure\ with\ PVI\ in\ all\ targeted\ veins\ at\ index}$$

- **Freedom from documented (symptomatic and asymptomatic) AF:** Kaplan-Meier estimates and plots will be used to characterize the time to first documented (symptomatic and asymptomatic) AF recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of non-study catheter for PV isolation, and failure to have PFA delivery with the study catheter due to IRE system malfunctions) will also be considered a failure. Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis. The probability of freedom from documented (symptomatic and asymptomatic) AF at each follow-up timepoint post blanking will be presented.
- **Freedom from documented (symptomatic and asymptomatic) atrial arrhythmia (AF, AT or AFL):** Kaplan-Meier estimates and plots will be used to characterize the time to first documented (symptomatic and asymptomatic) AF/AT or AFL recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis. The probability of freedom from documented (symptomatic and asymptomatic) AF/AT or AFL at each follow-up timepoint post blanking will be presented. The following criteria will also be deemed failures (see Appendix 1 of protocol for definitions):
    - Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of non-study catheter for PV isolation, and failure to have PFA delivery with the study catheter due to IRE system malfunctions)
    - Taking a new class I/III AAD for AF/AT or AFL for a previously failed class I/III AAD at a greater than the highest ineffective historical dose for AF/AT or AFL during the effectiveness evaluation period
    - Greater than 2 repeat ablations for AF/AT or AFL in the blanking period or any repeat ablation for AF/AT or AFL during the effectiveness evaluation period

7.6.4 Neurological Evaluations

The following analyses of neurological evaluations will be conducted using the mITT analysis set and summarized descriptively. In addition to the timepoints explicitly stated below, subjects will undergo full neurological follow-up only if neurologic symptoms and/or cerebral ischemic lesions are identified in a prior evaluation; results for these additional neurological evaluations will be summarized descriptively. Neurological evaluation questionnaires, including Mini Mental State Examination (MMSE), National Institute of Health Stroke Scale (NIHSS), and Modified Rankin Scale (mRS) will only be used in countries with validated languages. Only subjects in countries where the neurological evaluation questionnaires are applied will be included in the associated analyses.

- **Neurological Exam:** The occurrence of new or worsening neurological deficits post-ablation compared to pre-ablation will be summarized descriptively by timepoint.
- **Cerebral Emboli:** The frequency, anatomical location (side and area), and size (diameter and volume) of asymptomatic and symptomatic cerebral emboli observed pre-ablation and new emboli observed post-ablation as determined by Magnetic Resonance Imaging (MRI) evaluations assessed by a central core lab will be summarized descriptively by timepoint.
- **Mini Mental State Examination (MMSE):** The MMSE includes tests of registration, orientation (time and place), recall, attention and calculation, and language and praxis (naming, repetition, comprehension, reading, writing, and drawing)[2]. The total score ranges from 0 to 30, with lower scores indicating greater cognitive impairment. MMSE scores pre-ablation and change from pre-ablation at the 1-month follow-up will be summarized descriptively. MMSE scores will also be plotted by timepoint.
- **National Institute of Health Stroke Scale (NIHSS):** The NIHSS is a 15-item quantitative measure of stroke-related neurologic deficit on the levels of consciousness, language, neglect, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss[3]. The total score ranges from 0 to 42, with scores equal to 0 indicating no stroke symptoms, score ranging 1-4 indicating minor stroke, ranging 5-15 a medium stroke, 16-20 a moderate to severe stroke, and a range 21-42 a severe stroke. NIHSS scores pre-ablation and post-ablation prior to discharge will be summarized descriptively and plotted by timepoint.
- **Modified Rankin Scale (mRS):** The mRS is a disability scale with possible scores ranging from 0 to 6, where 0 indicates no symptoms and 6 indicates subject expiration[4]. mRS scores pre-ablation and change from pre-ablation at the 1-month follow-up will be summarized descriptively. mRS scores will also be plotted by timepoint.

7.6.5 CT/MRA

Occurrence of severe PV stenosis: Severe PV stenosis is defined as a 70% or more reduction in PV diameter compared to the PV diameter measured at baseline. The number and percentage of subjects and number and percent of targeted PVs having severe PV stenosis at 3 months post-ablation as assessed by a central core lab will be summarized in the mITT analysis set.

7.6.6 Endoscopy

Occurrence of esophageal lesions: The number and percentage of subjects experiencing esophageal lesions as determined by post-procedure endoscopy assessed by a central core lab will be summarized in the mITT analysis set.

8 DATA MONITORING COMMITTEE

A Data Monitoring Committee (DMC) will assess subjects' data for safety on regular intervals for Wave I and Wave II and make recommendations on study adaptations as described in the DMC Charter. There will be no formal interim analysis (sample size analysis or early success analysis).

9 REFERENCES

1. Spertus, J., et al., *Development and validation of the Atrial Fibrillation Effect on QualiTy-of-Life (AFEQT) Questionnaire in patients with atrial fibrillation.* Circ Arrhythm Electrophysiol, 2011. **4**(1): p. 15-25.
2. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state": A practical method for grading the cognitive state of patients for the clinician.* Journal of Psychiatric Research, 1975. **12**(3): p. 189-198.
3. Schlegel, D., et al., *Utility of the NIH Stroke Scale as a predictor of hospital disposition.* Stroke, 2003. **34**(1): p. 134-7.
4. Wilson, J.T., et al., *Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin Scale.* Stroke, 2002. **33**(9): p. 2243-6.

Statistical Analysis Plan

inspIRE (Wave II)

Pulsed Field Ablation (PFA) System for the Treatment of Paroxysmal Atrial Fibrillation (PAF) by Irreversible Electroporation (IRE)

Protocol # BWI_2019_08 (v 1.0)

Version 1.0

20 January 2022

Sponsor: BIOSENSE WEBSTER, INC.
31 Technology Dr.
Irvine, CA 92618
USA

Contents

List of Abbreviations

AAD	Antiarrhythmic Drug
AF	Atrial Fibrillation
AFEQT	Atrial Fibrillation Effect on QualiTy-of-life
AFL	Atrial Flutter
AT	Atrial Tachycardia
CEC	Clinical Events Committee
CSR	Clinical Study Report
CT	Computed Tomography
DMC	Data Monitoring Committee
ECG	ElectroCardioGram
FAS	Full Analysis Set
HRQoL	Health Related Quality of Life
ICF	Informed Consent Form
IRE	Irreversible Electroporation
mITT	Modified Intent-To-Treat
MMSE	Mini Mental State Examination
MRA	Magnetic Resonance Angiogram
MRI	Magnetic Resonance Imaging
mRS	Modified Rankin Scale
NIHSS	National Institutes of Health Stroke Scale
NSC	Non-study catheter
PAE	Primary Adverse Event
PAF	Paroxysmal Atrial Fibrillation
PFA	Pulsed Field ablation
PP	Per-Protocol
PV	Pulmonary Vein

PVI	Pulmonary Vein Isolation
QOL	Quality of Life
RF	Radiofrequency
SADE	Serious Adverse Device Effects
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan

1 STUDY DESIGN

This clinical investigation is an interventional, prospective, single arm, multi-center, pre-market clinical evaluation of the IRE ablation system (Circular IRE Catheter and IRE Generator) to demonstrate safety and effectiveness when compared to predetermined performance goals. The study will enroll subjects with drug refractory, symptomatic paroxysmal atrial fibrillation (PAF) who are candidates for catheter ablation in two sequential waves. Wave I subjects will undergo the index procedure and additional neurological, pulmonary vein (PV), and esophageal assessments. Wave II will consist of roll-in and main study phase subjects. Subjects in Wave II will undergo the index procedure and same follow-up schedule as Wave I subjects, excluding the additional neurological, post-procedural PV (unless symptomatic), and esophageal assessments. This statistical analysis plan (SAP) describes the analyses that will be performed for roll-in and main phase subjects in Wave II. A separate SAP describes the analyses for subjects in Wave I.

To verify consistent workflow for study device components and to minimize the learning curve effect on the evaluation of safety and effectiveness of the IRE system, a maximum of 180 roll-in subjects (the first 1-3 subjects for each ablating physician) will be enrolled in the study. The size of the roll-in cohort may be smaller if enrollment in the main study phase terminates at the first sample size interim analysis. Operators who performed procedures in Wave I may be exempt from roll-ins for Wave II per training charter. Roll-in subjects will not be counted towards the enrollment cap of 330 main study subjects.

A Bayesian adaptive design will be used to determine sample size for Wave II. A maximum of 330 main study phase subjects is planned. The number of subjects at any individual site shall not exceed 20% of the total enrollment. All Wave II subjects will be evaluated at 7 days, 1, 3, 6 and 12 months following the index procedure. An additional interim analysis will be performed to claim early trial success.

2 TREATMENT ASSIGNMENT

All subjects will be treated with the IRE ablation system, whose components consist of

- Multi-channel IRE generator (D-1417-01-I)
- Multi-electrode circular IRE catheter (D-1412-01-SI)
- Related components and accessories

3 RANDOMIZATION AND BLINDING PROCEDURES

This is a non-randomized trial with all subjects receiving treatment with the IRE ablation system. Therefore, masking of treatment assignment for operators and subjects will not be performed.

This study will employ several measures to minimize operational bias:

- Screening logs will be maintained at sites to confirm consecutive eligible subjects are considered for participation in the study
- Timing of the interim analyses for sample size selection will not be revealed to sites
- An independent statistician will be responsible for performing interim analyses
- Results from the interim analyses will not be shared with the Sponsor or sites unless the interim analysis results in a decision to stop enrollment or to file for approval
- Sponsor personnel directly involved in the conduct of the study will not have access to intermediate aggregated summaries of primary safety and effectiveness endpoint data until preparation for filing for approval

4 LEVELS OF SIGNIFICANCE

Significance levels for the interim and final analyses of the primary safety and effectiveness endpoints are described below in the relevant sections. The overall type-I error for the interim and final analyses of the primary safety and effectiveness endpoints are each controlled at a one-sided 2.5% level.

5 ANALYSIS SETS

5.1 Main Phase

For the analysis of study endpoints, the analysis sets defined in the following will be used:

- **Full Analysis Set (FAS):** The FAS will consist of all enrolled subjects who have had insertion of the study catheter (with or without delivery of pulse field ablation [PFA]).

- **Modified Intent-To-Treat (mITT) Analysis Set**: The mITT analysis set will consist of enrolled subjects who have had insertion of the study catheter and meet eligibility criteria.

- **Per Protocol (PP) Analysis Set**: The PP analysis set will consist of all enrolled subjects who satisfy the following criteria:
    - Have undergone ablation with the IRE ablation system.
    - Are treated for the study-related arrhythmia.
    - Are without major protocol deviations that would affect the scientific integrity of the primary safety and effectiveness data, including but not limited to the following:
        - Subjects found not meeting eligibility criteria
        - Use of the IRE ablation modality to ablate outside the PV region
        - Failure to check for entrance block for each targeted PV after adenosine/isoproterenol challenge

Eligible subjects treated with commercially available catheters for non-PV triggers that may arise during the index procedure or during repeat procedures occurring during the blanking period will be followed for the full duration of follow-up and will be evaluable for all endpoints in order to preclude potential bias that would arise from exclusion of these subject profiles from endpoint estimates.

5.2 Roll-in

- **Roll-In Analysis Set:** The Roll-In analysis set will include all subjects who are enrolled in the roll-in phase and have undergone ablation with the IRE ablation system.

6 SAMPLE SIZE JUSTIFICATION

A Bayesian adaptive design will be utilized to select the final sample size of the trial. Two sample size selection interim analyses will be performed based on the primary effectiveness endpoint. The first sample size selection interim analysis will be performed when enrollment in the mITT analysis set reaches 180. From simulations with realistic enrollment scenario, at this first interim timing, it is expected that about 30% of subjects will have between 3- and 6-months follow-up, about 18% of subjects will have between 6- and 9-months follow-up, and about 2% of subjects will have between 9- and 12-months follow-up. Predictive probabilities for trial success for the effectiveness endpoint will be used to determine whether the sample size of 180 will be sufficient or if the trial will continue enrollment. If the predictive probability with the current sample size of 180 is greater than 90%, or if the predictive probability with the maximum sample size of 330 is less than 2.5%, then enrollment will be stopped.

In case of continuation of enrollment, the second sample size interim analysis will be performed when enrollment in the mITT analysis set reaches 255. From simulations with realistic enrollment scenario, at this second interim timing, it is expected that about 27% of subjects will have between 3- and 6-months follow-up, about 27% of subjects will have between 6- and 9-months follow-up, and about 19% of subjects will have between 9- and 12-months follow-up. Predictive probabilities for trial success for the effectiveness endpoint will be used to determine whether the sample size of 255 will be sufficient or if the trial will continue to the full sample of 330. If the predictive probability with the current sample size of 255 is greater than 90%, or if the predictive probability with the maximum sample size of 330 is less than 2.5%, then enrollment will be stopped.

Under the assumption of 7% and 65% rates for the primary safety and effectiveness endpoints respectively, the study will be adequately powered for trial success (over 80%). The highest type I error rates were estimated in the following two scenarios: 1) the true primary safety rate was on the decision boundary (i.e., equal to 14%) and the device was assumed to be very effective, and 2) the true primary effectiveness rate was on the decision boundary (i.e., equal to 50%) and the device was assumed to be very safe. In both scenarios, the overall type I error for claiming success for both safety and effectiveness is controlled at 2.5%. For details of the simulation results of power and type I error, please refer to Appendix I - Adaptive Design Simulation Report.

7 STATISTICAL ANALYSIS METHODS

7.1 General Conventions

Standard descriptive summaries for continuous data include the number of observations with data, mean, standard deviation, median, 25th percentile, 75th percentile, minimum, and maximum values. For categorical data, the count and percent will be provided. Percentages will be based on the number of subjects without missing data.

7.2 Subject Disposition

Disposition and accountability of the study subjects will be summarized descriptively for the following subject categories:

- **Enrolled Subjects:** Patients who sign the informed consent form (ICF).
- **Excluded Subjects:** Subjects who are enrolled but never undergo insertion of the study catheter. Excluded subjects will be subjected to safety event reporting between ICF signature and date of exclusion. Subjects who signed the ICF but are found to be ineligible prior to insertion of the catheter are also considered as excluded.
- **Discontinued Subjects**: Subjects in whom the catheter is inserted but do not undergo ablation (i.e., no energy is delivered with the study catheter).
    - Discontinued subjects will remain in follow-up for 3-months post catheter insertion.
    - If an SAE is reported for a discontinued subject, the subject will be followed until event resolution (with or without sequelae), stabilization, or until the event is adequately explained.
- **Lost to Follow-up Subjects:** Subjects in whom the catheter is inserted for which contact is lost after most recent visit (despite 3 documented attempts to contact the subject).
- **Withdrawn / Early Termination Subjects:** Subjects who withdraw consent for study participation or are withdrawn by the investigator, are terminated from the study prior to completion of all follow-up visits.
- **Completed Subjects:** Enrolled subjects who have not been excluded, discontinued, withdrawn, terminated early, or lost-to-follow-up from the study prior to the final study visit.

7.3 Demographic and Baseline Characteristics

Subject demographics, medical history, previously failed antiarrhythmic drugs (AADs), active AAD use at baseline, and other baseline data will be summarized descriptively for subjects in the FAS, mITT, PP, and Roll-in analysis sets.

7.4 Analysis of Primary Endpoints

A single interim analysis is planned for an early success evaluation. The interim analysis will be performed when 30 subjects complete full 12 months of follow-up and all subjects complete 3 months of follow-up in the mITT analysis set. If early success is achieved, an early success interim analysis Clinical Study Report (CSR) will be compiled to present the results of the study success. If early success is not achieved, then the final primary endpoint analyses will be performed at the end of the study using the full follow-up data.

7.4.1 Primary Safety Endpoint

The primary safety endpoint is the occurrence of any PAE within 7 days of the initial mapping and ablation procedure. The definition of PAE can be found in section 8.2.1 of the study protocol.

The PAE rate will be compared against a performance goal of 14% by testing the following hypotheses:

$$H_0: p_s \geq 0.14 \qquad vs \qquad H_A: p_s < 0.14\,,$$

Where p_s is the PAE rate. The mITT analysis set will be the primary analysis set for the early success interim analysis and final analysis of primary safety endpoint.

7.4.1.1 Early Success Interim Analysis of Primary Safety Endpoint

The primary safety outcome for all subjects will be known at the time of the interim analysis and will be used for hypothesis testing. Notably, primary adverse events only occur within a 3-month window of the procedure by definition (with the exception of a procedure or device related death, pulmonary vein stenosis, and atrio-esophageal fistula) which will result in nearly identical hypothesis tests at the interim and final analyses.

At the time of the early success interim analysis, the primary safety endpoint will be met if:

$$\mathbf{Pr}(\boldsymbol{p_S < 0.14 | y, n}) > \mathbf{0.975}$$

where p_s is the PAE rate. A beta-binomial model with a non-informative uniform prior Beta (1,1) will be used for the safety rate. If the posterior probability of the safety rate being less than 14% is greater than 0.975 then the study will be considered to have demonstrated safety of the device at the early success interim analysis.

To investigate the robustness of the analysis result, the following sensitivity analyses will be performed in the FAS and mITT analysis sets unless otherwise noted at the time of the early success interim analysis with available data:

- **Sensitivity to Analysis Set**

  The primary safety endpoint will be analyzed in the FAS as a sensitivity analysis.

- **Snap-Shot Binomial Analysis**

  The number and percentage of subjects with primary safety events will be presented along with one-sided 97.5% exact binomial upper confidence bounds. Subjects with missing outcomes will be excluded from the analysis. Missing outcomes of primary safety endpoints is defined in Section 7.4.3.

- **Best-case Scenario**

  The point estimate for the PAE rate will be estimated by treating subjects with missing primary safety outcomes as free from primary safety events.

- **Worst-case Scenario**

  The point estimate for the PAE rate will be estimated by treating subjects with missing primary safety outcomes as failures.

- **Tipping Point Analysis**

  Tipping point analysis will be performed for the primary safety endpoint to assess the impact of missing outcomes on the safety conclusion. The posterior distribution will be updated each time treating a subject with a missing outcome as failure to evaluate whether a tipping point is identified.

7.4.1.2 Final Analysis of Primary Safety Endpoint

If early success is not achieved, then the final primary safety endpoint analysis will be performed at the end of the study.

The final analysis for the primary safety endpoint will use a beta-binomial model with a non-informative uniform prior Beta (1,1). At the time of the final analysis, the primary safety endpoint will be considered as success if:

$$\Pr(p_S < 0.14 | y, n) > 0.975$$

where p_s is the PAE rate. If the posterior probability of the safety rate being less than 14% is greater than 0.975 then the study will be considered to have demonstrated safety of the device.

To investigate the robustness of the analysis result, the following sensitivity analyses will be performed in the FAS and mITT analysis sets unless otherwise noted at the time of the final safety analysis:

- **Sensitivity to Analysis Set**

  The primary safety endpoint will be analyzed in the FAS as a sensitivity analysis.

- **Snap-Shot Binomial Analysis**

  The number and percentage of subjects with primary safety events will be presented along with one-sided 97.5% exact binomial upper confidence bounds. Subjects with missing outcomes will be excluded from the analysis. Missing outcomes of primary safety endpoints is defined in Section 7.4.3.

- **Best-case Scenario**

  The point estimate for the PAE rate will be estimated by treating subjects with missing primary safety outcomes as free from primary safety events.

- **Worst-case Scenario**

  The point estimate for the PAE rate will be estimated by treating subjects with missing primary safety outcomes as failures.

- **Tipping Point Analysis**

  Tipping point analysis will be performed for the primary safety endpoint to assess the impact of missing outcomes on the safety conclusion. The posterior distribution will be updated each time treating a subject with a missing outcome as failure to evaluate whether a tipping point is identified.

If early success is achieved, the primary safety endpoint and associated sensitivity analyses based on full 12-month follow-up data will be summarized descriptively.

7.4.2 Primary Effectiveness Endpoint

The primary effectiveness endpoint is defined as freedom from documented (symptomatic and asymptomatic) atrial arrhythmia (atrial fibrillation [AF], atrial tachycardia [AT] or atrial flutter

[AFL]) recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of a non-study catheter for PV isolation, or failure to have PFA delivery with the study catheter due to IRE system malfunctions) will also be considered a long-term effectiveness failure. Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis. The primary analysis of the primary effectiveness endpoint will be performed in the PP analysis set.

The primary effectiveness endpoint will be assessed by testing the following hypothesis:

$$H_0: p_E \leq 0.50 \qquad vs \qquad H_A: p_E > 0.50\,,$$

Where p_E is the effectiveness success rate.

7.4.2.1 Early Success Interim Analysis of Primary Effectiveness Endpoint

At the time of the early success interim analysis, the primary effectiveness endpoint will be met if:

$$\mathbf{Pr}(\boldsymbol{p_E} > \mathbf{0.50}|\boldsymbol{x}, \boldsymbol{n}) > \mathbf{0.9975}$$

where p_E is the effectiveness success rate. If the posterior probability of the effectiveness success rate being greater than 50% is greater than 0.9975 then the study will be considered to have demonstrated effectiveness of the device at the early success interim analysis.

At the time of the early success interim analysis, the time-to-failure during the 9-month (39 weeks) post-blanking period will be modeled by using a piecewise exponential model. The model will consist of three distinct time intervals during the 9-month evaluation period: (0, 2], (2, 8], and (8, 39] weeks. Failure times during each time interval is assumed to be exponentially distributed with a different hazard rate in each segment. The model is:

$$f(t) = \exp(-th(t))$$

where

$$h(t) = \begin{cases} \lambda_1 & 0 < t \leq 2 \\ \lambda_2 & 2 < t \leq 8 \\ \lambda_3 & 8 < t \leq 39 \end{cases}$$

Vague Gamma prior distributions will be assumed for λ_1, λ_2, and λ_3. Given hazard rates λ_1, λ_2, and λ_3, the failure-free rate at 12 months (9 months post-blanking) p_E can then be estimated

by

$$p_E = \exp(-[2\lambda_1 + (8-2)\lambda_2 + (39-8)\lambda_3])$$

At the interim, the total number of primary effectiveness events and the total subject exposure time will be used to update the vague Gamma prior and obtain the posterior Gamma distribution for each segment. Then 10,000 random hazard rates will be sampled from the posterior distribution for each segment and each triplet set of hazard rates will be used to calculate p_E. The distribution of p_E is estimated based on the 10,000 calculated $p_E s$. If the probability of p_E being greater than 50% is greater than 0.9975 (99.75%), the trial will declare early success.

To investigate the robustness of the analysis result, the following sensitivity analyses will be performed in the mITT and PP analysis sets at the time of the early success interim analysis with the available data:

- **Sensitivity to Analysis Set**

  The primary effectiveness endpoint will be analyzed in the mITT analysis set as a sensitivity analysis.

- **Kaplan-Meier Analysis**

  Kaplan-Meier estimates and plots will be used to characterize the time to first primary effectiveness event, including acute procedural failure and documented symptomatic/asymptomatic AF, AT or AFL recurrence following the initial ablation procedure. This will provide a non-parametric estimate of the primary effectiveness endpoint. The Kaplan-Meier analysis will account for all follow-up time for all subjects in the analysis set. Subjects with incomplete follow-up who remain failure free will be censored at the time of last observation. The probability of freedom from primary effectiveness failure at each monthly timepoint post blanking will be presented along with the one-sided 99.75% lower confidence bound using Greenwood's variance.

7.4.2.2 Final Analysis of Primary Effectiveness Endpoint

If trial success is not declared at the early success interim, the final analysis for the primary effectiveness endpoint will use a beta-binomial model with a non-informative uniform prior Beta (1,1) based on complete 12-month follow-up data. At the time of the final analysis, the primary effectiveness endpoint will be considered as success if:

$$\Pr(p_E > 0.50|x, n) > 0.9775$$

where p_E is the 12-month effectiveness success rate. If the posterior probability of the effectiveness rate being greater than 50% is greater than 0.9775 then the study will be considered to have demonstrated effectiveness of the device.

To investigate the robustness of the analysis result, the following sensitivity analyses will be performed in the mITT and PP analysis sets unless otherwise noted at the time of the final effectiveness analysis:

- **Sensitivity to Analysis Set**

  The primary effectiveness endpoint will be analyzed in the mITT analysis set as a sensitivity analysis.

- **Snap-Shot Binomial Analysis**

  Number and percentage of subjects free from primary effectiveness failure will be presented along with one-sided 97.75% exact binomial lower confidence bounds. Subjects with missing outcomes will be excluded from this analysis. Missing outcomes of primary effectiveness endpoints is defined in Section 7.4.3.

- **Kaplan-Meier Analysis**

  Kaplan-Meier estimates and plots will be used to characterize the time to first primary effectiveness event, including acute procedural failure and documented symptomatic/asymptomatic AF, AT, or AFL recurrence following the initial ablation procedure. This will provide a non-parametric estimate of the primary effectiveness endpoint. The Kaplan-Meier analysis will account for all follow-up time for all subjects in the analysis set. Subjects with incomplete follow-up who remain failure free will be censored at the time of last observation. The probability of freedom from primary effectiveness failure at each monthly timepoint post blanking will be presented along with the one-sided 97.75% lower confidence bound using Greenwood's variance.

- **Best-case Scenario**

  The point estimate for freedom from primary effectiveness failure will be estimated by treating subjects with missing primary effectiveness outcomes as free from primary effectiveness events.

- **Worst-case Scenario**

The point estimate for freedom from primary effectiveness failure will be estimated by treating subjects with missing primary effectiveness outcomes as failures.

- **Tipping Point Analysis**

  Tipping point analysis will be performed for the primary effectiveness endpoint to assess the impact of missing outcomes on the effectiveness conclusion. The posterior distribution will be updated each time treating a subject with a missing outcome as failure to evaluate whether a tipping point is identified.

If early success is achieved, the primary effectiveness endpoints and associated sensitivity analyses based on full 12-month follow-up data will be summarized descriptively – raw rates based on completers, and Kaplan-Meier estimates based on full follow-up will be presented.

7.4.3 Handling of Missing Data

Missing data will be queried for reasons and handled on an individual basis.

- **Sample Size Selection Interim Analysis:** For the sample size selection interim analyses, subjects with incomplete follow-up data for the effectiveness endpoint will be censored at the time of their last follow-up visit, and only their observed partial follow-up time will contribute to the estimation of model parameters and compute the predictive probability of the effectiveness success.

- **Early Success Primary Safety Endpoint:** For the early success interim analysis of the safety endpoints, if a subject's 3-month follow-up is not complete but a PAE has occurred prior to 3-month follow-up, the subject will be considered as having an event. If a subject exits the study without completing their 3-month follow-up visit and the subject has not had a PAE, that subject will be excluded from the early success safety analysis.

- **Early Success Primary Effectiveness Endpoint:** For the early success interim analysis of the effectiveness endpoints, subjects with incomplete follow-up data will be censored at the time of their last follow-up visit (if event-free). For these subjects, only their observed partial follow-up time will contribute to the estimation of model parameters. Subjects who refuse the remote arrythmia/holter monitoring device will be excluded from the early success effectiveness analysis.

- **Final Primary Safety Endpoint:** For the final primary safety analysis, if a subject does not have full 12 months of follow-up and the subject has had a PAE at any time during the follow-up, the subject will be considered an event. If a subject does not have 12 months of follow-up but has at least 3 months of follow-up (similar to early success interim), this subject's last observed status will be carried forward for the 12-month safety endpoint as a proxy. If a subject without a PAE does not have at least 3 months of follow-up, this subject will be considered as missing the primary safety endpoint and will be excluded from the final primary safety analysis.

- **Final Primary Effectiveness Endpoint:** For the final effectiveness analysis, if a subject has an effectiveness failure at any time during the evaluation period, then the subject will be considered an event. Subjects without an effectiveness failure who (1) do not have full 12 months of follow-up and/or sufficient follow-up duration for the primary effectiveness endpoint (i.e., with at least 335 days of follow-up or arrythmia monitoring post the index procedure), or (2) are missing all remote arrythmia and holter monitoring records as required by protocol will be considered missing primary effectiveness endpoint and will be excluded from the final primary effectiveness analysis.

7.4.4 Subgroup Analyses

In order to provide additional characterization and interpretation of the primary effectiveness and safety outcomes, the following subgroup analyses may be performed at the time of early success and final analysis in the mITT and PP analysis sets for the primary safety and effectiveness endpoints, respectively. Descriptive statistics will be presented for the primary safety and effectiveness outcomes in each subgroup. No inferential statistical analysis will be performed to test for differences between subgroups. Confidence intervals may be constructed for clinical relevance only. Descriptive statistics will consist of the number and rate of primary adverse events for the primary safety endpoint and Kaplan-Meier plots and estimates for the primary effectiveness endpoint in each subgroup.

- **Baseline Characteristics:**
  - Age group: <65 vs. ≥65 years
  - Sex
  - CHA_2DS_2-VASC Score: ≤2 vs. >2

7.4.5 Exploratory Analysis

Univariate and multivariate logistic regression analyses will be performed to examine the impact of demographic, baseline, and relevant procedural characteristics (e.g., age, sex, LA size, AF duration, medical history, number of PFA applications, etc.) on the primary outcomes. The outcome variables (i.e., primary effectiveness and safety outcomes) will be the dependent variables and baseline, demographic, and procedural characteristics will be treated as independent variables. Univariate logistic regression models will be conducted for each of the potential predictors. A p-value < 0.20 will be used as the cut-off for screening covariates. A multivariable model will be fit including the selected covariates from the univariate models. A final parsimonious multivariable model will be constructed by taking into consideration multicollinearity and clinical relevance of variables that retain a p-value <0.20 when modelled jointly. The identification of relevant associations with baseline factors provides additional characterization and interpretation of the primary effectiveness and safety outcomes.

7.4.6 Site Heterogeneity

Descriptive summaries of the primary safety and effectiveness endpoints will be provided by site. The PAE and primary effectiveness success rates will be presented by sites in the mITT and PP analysis sets, respectively. This analysis will be performed at the final analysis when all subjects complete their 12-month follow-up. If early success is declared, PAE estimates by site will be presented for the primary safety endpoint. As follow-up will be ongoing at the time of early success if declared, no site level estimates will be presented for the primary effectiveness endpoint.

7.5 Analysis of Secondary Endpoints

No formal statistical hypothesis or inferential statistics will be formulated or performed for the secondary endpoints. Descriptive statistics will be provided on all secondary endpoints in the mITT and PP analysis sets, unless otherwise specified. Acute procedural success rates and Kaplan-Meier estimates and plots for freedom from documented symptomatic recurrence will be presented at the time of early success interim analysis with available data. All secondary endpoints will be analyzed at the time of final analysis using full 12-month follow-up data.

7.5.1 Acute Procedural Success

Acute procedural success is defined as confirmation of entrance block in all clinically relevant targeted PVs after adenosine/isoproterenol challenge. Use of a non-study catheter for PVI and

failure to have PFA delivery with the study catheter due to IRE system malfunctions are considered acute procedural failures.

The rate of acute procedural success will be summarized descriptively. The number and percentage of subjects with acute procedural success will be presented.

7.5.2 Freedom from Documented Symptomatic Recurrence

The symptomatic recurrence endpoint is defined as freedom from documented symptomatic atrial arrhythmia (AF, AT or AFL) recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of non-study catheter for PVI, and failure to have PFA delivery with the study catheter due to IRE system malfunctions) will also be considered a failure. Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis.

Symptomatic recurrence will be summarized descriptively as the number and percent of subjects event-free during the evaluation period. Kaplan-Meier estimates and plots will be used to characterize the time to first documented symptomatic AF, AT or AFL recurrence following the initial ablation procedure. The probability of freedom from symptomatic AF, AT or AFL recurrence at each follow-up timepoint post blanking will be presented.

7.5.3 Quality of Life (QOL)

The Atrial Fibrillation Effect on Quality-of-Life (AFEQT™) includes 20 questions on a 7-point Likert scale. Questions 1-18 evaluate Health Related Quality of Life (HRQoL) and questions 19-20 relate to patients' satisfaction with treatment. The first 18 questions are used to calculate the Overall AFEQT score and the subscale scores across three domains as follows:

- AF related Symptoms: Four questions (1 – 4)
- Daily Activities: Eight questions (5 – 12)
- Treatment Concerns: Six questions (13 – 18)

The formula to calculate these scores is as follows:

$$100 - [\frac{(sum\ of\ severity\ for\ all\ questions\ answered - number\ of\ questions\ answered)\ \times 100}{total\ number\ questions\ answered \times 6}]$$

Overall and subscale scores range from 0 to 100. A score of 0 corresponds to complete disability, while a score of 100 corresponds to no disability.

Baseline AFEQT scores and changes from baseline at each timepoint the questionnaire is administered will be summarized descriptively for the following five scores. The overall AFEQT score and subscale scores across study visits will also be plotted.

- Overall AFEQT Score (18 questions)
    - Symptom Subscale Score (4 questions)
    - Daily Activities Subscale Score (8 questions)
    - Treatment Concern Subscale Score (6 questions)
- Treatment Satisfaction Score (2 questions)

The AFEQT questionnaire will only be used in countries with validated languages. These analyses will be conducted for subjects in countries where the AFEQT questionnaires are applied.

7.6 Analysis of Additional Endpoints

No formal statistical hypothesis or inferential statistics will be formulated or performed for the additional endpoints. Descriptive statistics will be provided on all additional endpoints in the analysis sets specified below. Procedural data, additional safety endpoints, and additional effectiveness endpoints using Kaplan-Meier estimates and plots where relevant will be presented at the time of early success interim analysis with available data. All additional endpoints will be analyzed at the time of final analysis using full 12-month follow-up data.

7.6.1 Procedural data

Procedural data such as total procedure time, mapping time, PFA application time, number of PFA applications by PV and by subject, total fluoroscopy time, study catheter dwell time, ablation settings used, and use of anesthesia will be summarized descriptively for subjects in the PP analysis set.

7.6.2 Additional Safety Endpoints

The following analyses for the additional safety endpoints will be conducted in the FAS and mITT analysis sets and summarized descriptively as the total number of events, number of subjects with events, and percentage of subjects with events:

- Occurrence of individual PAEs from the primary composite
- Occurrence of Serious Adverse Device Effects (SADEs)

- Occurrence of Serious Adverse Events (SAEs) within 7 days (early-onset), 8-30 days (peri-procedural), and >30 days (late onset) of initial ablation procedure, separately for each timeframe
- Occurrence of non-SAEs
- Occurrence of clinically symptomatic severe PV stenosis as documented by computed tomography (CT)/ magnetic resonance angiogram (MRA) as assessed by a central core lab

7.6.3 Additional Effectiveness Endpoints

The following analyses for the additional effectiveness endpoints will be conducted in the mITT and PP analysis sets unless otherwise specified and summarized descriptively:

- **Ablation by non-study catheter for PV isolation**: The number and percentage of subjects and number and percent of clinically relevant targeted PVs ablated by a non-study catheter (NSC) for PVI will be summarized.
    - Rate of ablation by NSC for PVI among all clinically relevant targeted pulmonary veins:
    $$= \frac{Number\ of\ PVs\ ablated\ by\ NSC}{Number\ of\ ablated\ PVs}$$
    - Rate of ablation by NSC for PVI among subjects:
    $$= \frac{Number\ of\ subjects\ \ with\ at\ least\ one\ PV\ ablated\ by\ NSC}{Total\ number\ of\ subjects\ ablated\ for\ PVI}$$

- **Acute PV reconnection:** The number and percentage of subjects and number and percentage of targeted PVs where acute reconnection was identified by adenosine/isoproterenol challenge.
    - Rate of acute PV reconnection among targeted veins:
    $$= \frac{Number\ of\ PVs\ with\ reconnection\ after\ adenosine/isoproterenol\ challenge}{Total\ number\ of\ PVs\ undergoing\ adenosine/isoproterenol\ challenge}$$
    - Rate of acute PV reconnection among subjects:

$$= \frac{Number\ of\ subjects\ with\ PV\ reconnection\ in\ at\ least\ one\ PV\ after\ adenosine/isoproterenol\ challenge}{Total\ number\ of\ subjects\ undergoing\ adenosine/isoproterenol\ challenge\ in\ all\ targeted\ PVs}$$

- **Repeat ablation during the 12-month follow-up period**
    - **Occurrence of repeat procedures:** Kaplan-Meier estimates and plots will be used to characterize the time to the first repeat ablation procedure for study related arrythmia. The probability of freedom from repeat ablation for study related arrythmia at each follow-up timepoint post index procedure will be presented. Additionally, the total number of repeat procedures, number of subjects undergoing a repeat procedure, and percentage of subjects undergoing a repeat procedure will be presented by type of arrhythmia treated (overall (all

arrhythmias), AF and AFL or AT (left atrium)), and by timing of the repeat procedure (during the blanking period (Day 1-90) and during the evaluation period (Day 91-365)).

- **PV reconnection:** The number and percentage of subjects and number and percentage of targeted PVs at the repeat procedure where PV reconnection was observed
  - Rate of PV reconnection among previously isolated veins:

$$= \frac{Number\ of\ previously\ isolated\ PVs\ with\ reconnection\ identified\ at\ repeat}{Total\ number\ of\ previously\ isolated\ veins\ in\ subjects\ undergoing\ repeat\ ablation}$$

  - Rate of PV reconnection among subjects:

$$= \frac{Number\ of\ subjects\ with\ PV\ reconnection\ identified\ at\ repeat\ in\ at\ least\ 1\ previously\ isolated\ vein}{Total\ number\ of\ subjects\ undergoing\ repeat\ ablation\ with\ PVI\ in\ all\ targeted\ veins\ at\ index}$$

- **Freedom from documented (symptomatic and asymptomatic) atrial fibrillation (AF):** Kaplan-Meier estimates and plots will be used to characterize the time to first documented (symptomatic and asymptomatic) AF recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of non-study catheter for PV isolation, and failure to have PFA delivery with the study catheter due to IRE system malfunctions) will also be considered a failure. Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis.

  The probability of freedom from documented (symptomatic and asymptomatic) AF at each follow-up timepoint post blanking will be presented. The event-free rate will also be summarized descriptively as the number and percent of subjects free from documented AF during the evaluation period.

- **Freedom from documented (symptomatic and asymptomatic) atrial arrhythmia (AF, AT or AFL) with additional failure modes:** Kaplan-Meier estimates and plots will be used to characterize the time to first documented (symptomatic and asymptomatic) AF, AT or AFL recurrence based on electrocardiographic data (≥30 seconds on arrhythmia monitoring device) during the effectiveness evaluation period (Day 91-Day 365). Subjects without effectiveness failure who undergo a repeat ablation procedure during the blanking period where PVs were ablated with a commercially available catheter and generator will be considered non-assessable for long-term effectiveness of the IRE ablation system and thus will be excluded from this analysis.

The probability of freedom from documented (symptomatic and asymptomatic) AF, AT or AFL at each follow-up timepoint post blanking will be presented. The event-free rate will also be summarized descriptively as the number and percent of subjects free from documented AF, AT, or AFL during the evaluation period. The following criteria will also be deemed failures:

- Acute procedural failure (i.e., failure to confirm entrance block in all PVs except those that are silent and/or cannot be cannulated post-procedure, use of non-study catheter for PV isolation, and failure to have PFA delivery with the study catheter due to IRE system malfunctions)
- Taking a new class I/III AAD for AF, AT or AFL or a previously failed class I/III AAD at a greater than the highest ineffective historical dose for AF, AT or AFL during the effectiveness evaluation period
- Greater than 2 repeat ablations for AF, AT or AFL in the blanking period or any repeat ablation for AF, AT or AFL during the effectiveness evaluation period

7.7 Analysis of Roll-In Subjects

Descriptive statistics will be presented for all Wave II endpoints for subjects in the Roll-in analysis set, separately.

8 DATA MONITORING COMMITTEE

A Data Monitoring Committee (DMC) will assess subjects' data for safety on regular intervals for Wave I and Wave II and make recommendations on study adaptations as described in the DMC Charter. An independent statistician will be responsible for conducting the interim analyses and reviewing the results with the designated DMC. The DMC charter will document the constitution, roles and responsibilities of the committee, sponsor, and the independent statistician.

9 REFERENCES

1. Spertus, J., et al., *Development and validation of the Atrial Fibrillation Effect on QualiTy-of-Life (AFEQT) Questionnaire in patients with atrial fibrillation.* Circ Arrhythm Electrophysiol, 2011. **4**(1): p. 15-25.
2. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state": A practical method for grading the cognitive state of patients for the clinician.* Journal of Psychiatric Research, 1975. **12**(3): p. 189-198.
3. Schlegel, D., et al., *Utility of the NIH Stroke Scale as a predictor of hospital disposition.* Stroke, 2003. **34**(1): p. 134-7.
4. Wilson, J.T., et al., *Improving the assessment of outcomes in stroke: use of a structured interview to assign grades on the modified Rankin Scale.* Stroke, 2002. **33**(9): p. 2243-6.

10 APPENDIX I – ADAPTIVE DESIGN SIMULATION REPORT

10.1 Study Design

This clinical investigation is a prospective, non-randomized, pre-market clinical evaluation of the IRE system (circular IRE catheter and IRE generator) to demonstrate safety and long-term effectiveness when compared to an historical performance goal. The sample size for the study is primarily driven by the effectiveness endpoint. A Bayesian adaptive design will be used to determine the sample size based on the effectiveness endpoint. A single Bayesian interim analysis for early success will be performed when 30 subjects full follow-up of one year and all subjects in the mITT analysis set complete 3 months of follow-up . The final analysis will use a beta-binomial model after all subjects complete the 12 months follow-up.

The effectiveness endpoint will be assessed by testing the hypotheses:

$$H_0: p_E \leq 0.50 \qquad \text{vs.} \qquad H_A: p_E > 0.50,$$

where p_E is the proportion of patients free from effectiveness events.

Let X be the number of patients that are failure-free through 12 months. We model the number of patients free from an event as

$$X \sim \text{Binomial}(n, p_E),$$

where n is the number of patients. We use a vague Beta(1, 1) prior distribution for p_E. Then the posterior distribution is

$$(p_E|x, n) \sim \text{Beta}(1 + x, 1 + n - x).$$

Similarly, the hypothesis test for the safety endpoint is:

$$H_0: p_S \geq 0.14 \qquad \text{vs.} \qquad H_A: p_S < 0.14,$$

where p_S is the rate of the Primary Adverse Events (PAEs). We model the number of patients with PAEs as

$$Y \sim \text{Binomial}(n, p_S)$$

with a vague Beta(1, 1) prior distribution on p_S.

The trial will be considered a success at the time of the final analysis if BOTH

1. $\Pr(p_E > 0.5|x, n) > 0.9775$ AND
2. $\Pr(p_S < 0.14|y, n) > 0.975.$

These thresholds control the overall Type I error rate for the trial below one-sided 2.5%.

10.2 Sample Size Justification

The methods described in Broglio et al.[1] will be used to determine the sample size based on the effectiveness endpoint. Sample size selection interim analyses will be performed when 180 and 255 patients are enrolled in the mITT analysis set. Predictive probabilities for trial success will be used to determine whether the sample size at the time of the interim analysis will be sufficient or

if the trial will continue to the full sample of 330. The power for testing the primary safety endpoint is greater than 80% at all sample sizes greater than or equal to180 subjects, assuming a primary safety rate of 7% with a performance goal of 14%.

It is assumed that there is a 5% dropout rate for the safety endpoint and a 15% dropout rate for the effectiveness endpoint. The sample size will provide greater than 80% power at one-sided significance level of 2.5% for each endpoint assuming that a true safety rate of 7% and a true effectiveness rate of 65% compared to the performance goals of 14% and 50% respectively.

10.3 Adaptive Sample Size Determination

Sample size selection interim analyses will be performed when 180 and 255 subjects (excluding roll-in) are enrolled in the mITT analysis set. Predictive probability of success for the effectiveness endpoint will be used to determine whether the sample size at each interim analysis will be sufficient or if the trial enrollment will continue. Sample size simulations were performed using performance goals of 14% and 50% respectively for the safety and effectiveness endpoints.

At the time of each interim analysis, predictive probability of success for the effectiveness will be calculated once using the current sample size and another time using the maximum sample size allowed. If the effectiveness predictive probability is greater than 90%, or if the effectiveness predictive probability with the maximum sample size of 330 is less than 2.5%, then the enrollment will be stopped. Otherwise the enrollment will continue.

10.3.1 Prediction Probability for the Effectiveness Endpoint

To estimate the predictive probability of success for the effectiveness endpoint at each interim analysis, time-to-failure during the 9-month (39 weeks) post-blanking period will be modeled by a piecewise exponential model with 3 distinct intervals (from (0, 2], (2, 8], and (8, 39] weeks). The three-piece exponential model provides a reasonable amount of flexibility in capturing the functional form of the survival function for the effectiveness endpoint.

A vague Gamma prior distribution will be assumed for estimating each of the hazard rates:

$$h_j \sim \Gamma(\alpha=0.1, \beta=0.1) \text{ for } j = 1,2,3.$$

where α represents the prior number of events and β the prior exposure time. At the sample size selection interim analysis with n enrolled patients, the total number of observed events for time interval j is EV_j and the total observed exposure time for time interval j is EXP_j, assuming each patient i offers EXP_{ij} exposure for time period j. For each time interval, we update the prior distribution with the currently observed data (EV_j, EXP_j) and the resulting posterior distribution is:

$$h_j \sim \Gamma(\alpha + EV_j,\ \beta + EXP_j).$$

The predictive probabilities of trial success will be calculated using Monte Carlo integration. For the predictive probability of trial success at the current sample size, we assume accrual is stopped and the primary analysis is conducted after all subjects complete the 12 months follow-up. The data will be imputed in the following fashion. First, a single hazard rate, h_j, is sampled from the posterior distribution of the hazard rates for each time interval. For each patient who has not yet experienced an event, depending on which time interval their current observed time falls in, we sample an event time from an exponential distribution with mean equal to the sampled hazard rate for that time interval*:*

$$TTE_{ij} \sim \text{Exponential}(h_j),$$

where TTE_{ij} represents the time to event for patient i in time period j. For patient i, his or her imputed exposure time for time interval j would be $EXP_{ij} + TTE_{ij}$. If the imputed exposure time is less than or equal to the end of the time interval, the patient will be considered an event, otherwise, the sampled $h_j$for the next time interval will be used to generate a time for the next time interval and impute the exposure time. The patient time will be censored at 39 weeks if the imputed time for the last time interval is greater than 39 weeks. Additionally, a censoring time will be sampled from an exponential distribution with a rate that results in the assumed attrition rate for the effectiveness endpoint at 1 year. If this censoring time is less than the event time for the subject, then the patient will be censored at this time and will not count as an event.

This process will be repeated 1000 times to generate 1000 imputed datasets. For each of these imputed datasets, success of the final analysis for the effectiveness endpoint will be assessed once assuming we would stop at the interim with the current sample size and another time assuming we use the maximum sample size of 330. The proportion of successful trials in each case provides the predictive probability of success for the effectiveness endpoint.

10.4 Interim Analyses for Early Success

One interim analysis is planned for an early success claim. The interim analysis will be performed when the first 30 subjects complete full follow-up of one year and all subjects in the mITT analysis set complete 3 months of follow-up. For the safety endpoint, the outcome for all patients will be known at the time of this interim analysis and the test will be final primary safety analysis.

At the interim analysis, the trial will declare early success if:

1. The safety objective is met.
2. Posterior probability of the effectiveness proportion p_E being greater than 50% is greater than 0.9975.

The posterior probability is estimated based on the proposed model in the next section.

10.5 Modeling for Effectiveness

For the effectiveness endpoint, the time-to-failure during the 9-month (39 weeks) post-blanking period is modeled by using a piecewise exponential model. The model has three distinct time intervals: (0, 2], (2, 8], and (8, 39] weeks. These intervals are based on a model fit to the ThermoCool Pivotal trial and the SmartTouch IDE trial. The three-piece exponential model

provides a reasonable amount of flexibility in capturing the functional form of the survival function for the effectiveness endpoint.

For the model-based approach, failure times during each interval is assumed to be exponentially distributed, with different hazard rates in each segment. The model is:

$$f(t) = \exp(-th(t)),$$

where

$$h(t) = \begin{cases} h_1 & 0 < t \leq 2; \\ h_2 & 2 < t \leq 8; \\ h_3 & 8 < t \leq 39. \end{cases}$$

A vague Gamma(1, 1) prior distribution will be used for each h in the model. Given hazard rates h_1, h_2 , and h_3 , the failure-free rate at 12 months p_E can then be estimated by

$$p_E = \exp(-[2h_1 + (8-2)h_2 + (39-8)h_3]).$$

At the time of the early success interim analysis, the total number of primary effectiveness events and the total subject exposure time will be used to update the vague Gamma prior and obtain the posterior Gamma distribution for each segment. Then 10,000 random hazard rates will be sampled from the posterior distribution for each segment and each triplet set of hazard rates will be used to calculate p_E. The distribution of p_E is estimated based on the 10,000 calculated $p_E s$. If the probability of p_E being greater than 50% is greater than 0.9975(99.75%), the trial will declare early success.

10.6 Power and Type-I Error Simulations

The operating characteristics of this trial were determined through trial simulation. We hypothesized several scenarios for the underlying rates for the effectiveness and safety endpoints, and simulated the entire trial multiple times under each scenario. In each virtual trial, the interim analysis was conducted according to the pre-specified rules, and results were tracked for each trial, including whether the trial was successful on each endpoint individually and on both endpoints, whether trial success was achieved early, etc. Three different enrollment schedules were considered for the simulations corresponding to realistic, slow, and fast enrollment schedules based on logistical considerations:

Table 1. Assumed enrollment rates

	Expected number of patients per month																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Realistic	7.5	10.5	13.5	7.5	15	18	24	30	30	30	30	30	30	30	24			
Slow	7.2	7.2	7.2	9.6	7.2	16.8	22.8	24	24	24	24	24	24	24	24	24	24	12
Fast	10.2	10.2	5.1	10.2	11.9	15.3	17	25.5	32.3	34	34	34	34	34	23.8			

10.6 1 Effectiveness Outcome

In order to simulate the effectiveness outcome for a virtual subject, we need to simulate whether the subject has a failure and when that failure occurs. As previously discussed in section 5, a piecewise exponential time-to-failure model is assumed and calibrated to have a particular failure rate $(1 - p_E)$ at the end of follow-up. The profiles are constructed by taking advantage of the relationship:

$$p_E = \exp(-[2h_1 + (8-2)h_2 + (39-8)h_3]),$$

where p_E is the failure-free rate at 12 months. Rates from a previous study were assumed and calibrated to simulate effectiveness outcomes. From ThermoCool SSED, the reported hazard rates have the following pattern:

$$h_1 = (38.06)h_3;$$

$$h_2 = (1.71)h_3;$$

$$h_3 = (1)h_3.$$

To simulate data, we used the above multiplicative factors from ThermoCool and then found the value for h_3, such that when multiplied by each of the assumed rates for the three time intervals, the failure-free rate at 12 months (39 weeks following 13 weeks blanking period) matches the desired scenarios. Table 2 shows the derived hazard rates that were used to simulate data for a range of assumed failure-free rates.

Table 2. Parameters to generate failure times

	Hazard Rate (*h*)		
Failure-free Rate	h_1	h_2	h_3
0.50	0.2248	0.0101	0.0059
0.60	0.1656	0.0074	0.0044
0.62	0.1550	0.0070	0.0041
0.65	0.1397	0.0063	0.0037
0.70	0.1157	0.0052	0.0030
0.99999	0.0001	0.0001	0.0001

10.6.2 Attrition Rate Profile

For each subject, a time-to-withdrawal is simulated from an exponential distribution with a 5% dropout rate for the safety outcome and a 15% dropout rate for the effectiveness outcome. Both dropout rates are per year of follow-up. If a subject experiences a failure prior to their withdrawal time, then their outcome is recorded as a failure. A subject is censored at the time of withdrawal if the simulated withdrawal time is prior to the simulated failure time.

10.6.3 Simulation Results

Table 3.a provides the estimated powers for various effectiveness rates assuming a true safety rate of 7% under the realistic enrollment scenario. The power for each effectiveness rate scenario is obtained based on 10,000 simulations. The table also provides the proportion of times we expect to stop enrollment at the interim look for adequate sample size, and the proportion of times early success is achieved. Assuming a true effectiveness rates of 62 to 70%, the power for the effectiveness endpoint is estimated to be greater than 80%. When we assume the true safety rate of 7%, the powers for the safety endpoint are over 85% across all effectiveness rate scenarios. Thus, the trial is adequately powered for various assumed effectiveness rates.

Table 3.a Estimated power for various effectiveness rates assuming a true safety rate of 7% using 10,000 simulations for each scenario

Safety Rate	Effectiveness Rate	Overall Power	Win Effect	Win Safety	N=180	N=255	N=330	Early Success
0.07	0.60	0.6749	0.7289	0.9149	0.3702	0.2191	0.4107	0.6202
	0.62	0.7764	0.8527	0.9026	0.4272	0.2653	0.3075	0.7436
	0.65	0.8395	0.9457	0.8853	0.5535	0.2854	0.1611	0.8316
	0.70	0.8550	0.9954	0.8588	0.7296	0.2357	0.0347	0.8547

Table 3.b shows Type I error rates for the trial under the worst-case scenarios using the realistic enrollment scenario; (1) Assuming the true effectiveness rate is on the decision boundary of 50% and the device is to be safe (the true safety rate of 0.001%); and (2) Assuming the true PAE rate is on the decision boundary of 14% and the device is to be effective (the true effectiveness rate is 99.999%). Since the primary safety analysis is performed at a 2.5% level Type-I error and the sample size is selected independent of the safety endpoint, the Type-I error for the safety hypothesis is controlled at 2.5% and is confirmed in Table 1.b (see the safety rate of 14% and effectiveness rate of 99.999%). For the worst-case scenario for the effectiveness endpoint, Table 1.b shows that the estimated Type-I error rate from the simulation trials was 1.68%, still below the 2.5% threshold.

Table 3.b Estimated Type-I error for the worst-case scenarios based on 25,000 simulations

Safety Rate	Effectiveness Rate	Type-I error	95% C.I
0.14	0.99999	0.0172	[0.0156 , 0.0189]
0.00001	0.50	0.0168	[0.0152 , 0.0185]

Table 3.c provides the proportion of times we would stop the trial at each interim for futility under the realistic enrollment scenario.

Table 3.c Proportion of times the trial will stop at each interim for futility on 10,000 simulations for each scenario

Safety Rate	Effectiveness Rate	Overall Futility	N=180	N=255
0.07	0.60	0.0034	0.0020	0.0014
	0.62	0.0012	0.0008	0.0004
	0.65	0.0003	0.0003	0.0000

	0.70	0.0001	0.0001	0.0000

Tables 4.a, 4.c, 5.a, and 5.c provide the estimated power and futility proportions under the slow and fast scenarios. Tables 4.b and 5.b provide the estimated type-I errors under the slow and fast scenarios.

Table 4.a Estimated power for slow enrollment assumption based on 10,000 simulations

Safety Rate	Effectiveness Rate	Overall Power	Win Effect	Win Safety	N=180	N=255	N=330	Early Success
0.07	0.60	0.6758	0.7377	0.9085	0.3768	0.2621	0.3611	0.6410
	0.62	0.7769	0.8624	0.8975	0.4711	0.3021	0.2268	0.7595
	0.65	0.8336	0.9469	0.8788	0.6111	0.2949	0.0940	0.8307
	0.70	0.8399	0.9956	0.8437	0.8121	0.1782	0.0097	0.8399

Table 4.b Estimated Type-I error for slow enrollment assumption based on 25,000 simulations

Safety Rate	Effectiveness Rate	Type-I error	95% C.I
0.14	0.99999	0.0182	[0.0166 , 0.0199]
0.00001	0.50	0.0167	[0.0151 , 0.0184]

Table 4.c Proportion of times the trial will stop at each interim for futility for slow enrollment assumption based on 10,000 simulations

Safety Rate	Effectiveness Rate	Overall Futility	N=180	N=255
0.07	0.60	0.0037	0.0023	0.0014
	0.62	0.0011	0.0009	0.0002
	0.65	0.0002	0.0002	0.0000
	0.70	0.0000	0.0000	0.0000

Table 5.a Estimated power for fast enrollment assumption based on 10,000 simulations

Safety Rate	Effectiveness Rate	Power	Win Effect	Win Safety	N=180	N=255	N=330	Early Success
0.07	0.60	0.6777	0.7284	0.9212	0.3584	0.2110	0.4306	0.5995
	0.62	0.7702	0.8439	0.9059	0.4316	0.2460	0.3224	0.7223
	0.65	0.8407	0.9459	0.8863	0.5641	0.2576	0.1783	0.8315
	0.70	0.8499	0.9953	0.8540	0.7604	0.1972	0.0424	0.8493

Table 5.b Estimated Type-I error for fast enrollment assumption based on 25,000 simulations

Safety Rate	Effectiveness Rate	Type-I error	95% C.I
0.14	0.99999	0.0172	[0.0156 , 0.0189]

0.00001	0.50	0.0159	[0.0144 , 0.0175]

Table 5.c Proportion of times the trial will stop at each interim for futility for fast enrollment assumption based on 10,000 simulations

Safety Rate	Effectiveness Rate	Overall Futility	N=180	N=255
0.07	0.60	0.0034	0.0024	0.0010
	0.62	0.0019	0.0014	0.0005
	0.65	0.0003	0.0003	0.0000
	0.70	0.0001	0.0001	0.0000

10.7 Power and Type-I Error Simulations

1.Broglio KR, Connor JT, Berry SM. (2014) Not too big, not too small: A Goldilocks approach to sample size selection. J Biopharmaceutical Statistics. 24(3):685-705.