Statistical Analysis Plan

Clinical Investigation Plan Title	SPYRAL HTN-ON MED STUDY
Clinical Investigation Plan Version	10.0
Document Version	5.5
Document Date	17-May-2022
NCT #	NCT02439775

Table of Contents

1. Version History

Version	Summary of Changes	Author(s)/Title
1.0	• First draft	Martin Fahy, Senior Principal Statistician
2.0	• Study populations updated. ABPM algorithm updated. Subgroup analyses updated	Martin Fahy, Senior Principal Statistician
3.0	• Updated to new template. Study populations updated	Martin Fahy, Senior Principal Statistician
4.0	• Powered SPYRAL HTN-ON MED Study	Martin Fahy, Senior Principal Statistician
4.1	• Bayesian design updated to remove test for expected success • Changes to discount function formulation	Martin Fahy, Senior Principal Statistician
4.2	• Per-protocol population updated to exclude subjects who did not receive their randomized treatment • Treatment effect for efficacy analysis updated so that a negative direction favors RDN • Stochastic comparison in section 9.9.4.1 updated • Prior distribution and hyper-parameter definitions in section 9.9.4.1 updated • Section 9.4 updated to include an analysis to impute missing outcome data • Subgroup analyses in section 9.11 updated • Appendices II, III and IV updated to include unscheduled visits corresponding to repeat ABPM visits	Martin Fahy, Senior Principal Statistician
4.3	• Updated language to reflect "approximately 149" subjects included in the first interim analysis due to variable attrition rate	Martin Fahy, Senior Principal Statistician
4.4	• Updated hyperparameter value for b_β to be 1e10, changed from 1e5	Martin Fahy, Senior Principal Statistician
5.0	• First and second interim analysis to take place with 95 and 149 randomized subjects. Sections 8, 9.8, 9.9.4.3, 9.9.5 updated.	Martin Fahy, Senior Principal Statistician

	• Section 9.12 added describing analyses related to COVID-19 • As Treated population added in 9.1.3.4 • Added "y_i is the BP change for the ith observation and x_i is the mean centered baseline BP for the ith observation" to section 9.9.4.1. for the definition of the x_i and y_i model parameters • Added "mean centered" to section 9.9.4.1. COMBINE for the definition of the x_i model parameter • Updated Table 3 in section 9.9.4.3 to include 2 decimal places for prior treatment effects • Interim analysis stopping rules in section 9.8 clafiried to state that enrollment will stop if stopping rules are met • Primary safety analysis section 9.9.1.1. updated to clarify that primary safety analysis will only be performed once using first 253 subjects • Primary Efficacy Endpoint Analysis section 9.9.4 updated to indicate that the primary analysis treatment effect is Bayesian posterior estimate • Section 9.9.5 updated to include frequentist ANCOVA analyses as secondary analysis of primary efficacy endpoints. Also specifies that secondary cohort subjects will be analyzed using frequentist ANCOVA and will be performed internally • Added attrition rate of 15% to Table 3 in section 9.9.4.3 • Subgroup analyses section updated to include same as OFFMED, to specify complements for every subgroup, and to indicate which subgroups analyses are restricted to the RDN arm • Appendix I (missing date imputation) updated • Appendix VII (section 11.7) added with sensitivy analyses of Bayesian operating characteristics	
5.1	• First and second interim analysis to take place with 130 and 175 randomized subjects. Sections 8, 9.8, 9.9.4.3, 9.9.5 updated • Bayesian sensitivity analyses added in section 9.9.5 • Appendices describing the algorithms for selecting office BP, 24-hour BP, lab values	Martin Fahy, Senior Principal Statistician

	and drug testing values removed from SAP, and will only be included in study table specificaitons • Validation requirements added in section 10. • SAP template version added to footer	
5.2	• Subgroup analyses for accessory arteries treated and 1 vs. 2 vs. 3 AH Medication Classes at baseline added in section 9.11	Martin Fahy, Senior Principal Statistician
5.3	• Win ratio analyses and reference added to section 9.9.9 • HTN Burden analyses and reference added to section 9.9.10 • % Time in Target Range (TTR) analyses and reference added in section 9.9.11	Martin Fahy, Senior Principal Statistician
5.4	• Three new Medication Adherence analysis populations added to section 9.1.3 • Witnessed pill intake and BP protocol deviations added to Additional Objectives in section 9.9.8	Martin Fahy, Senior Principal Statistician
5.5	• Win Ratio section 9.9.9 updated • TTR section 9.9.10 updated • HTN Burden section removed	Martin Fahy, Senior Principal Statistician

2. List of Abbreviations and Definitions of Terms

Abbreviation	Definition
ABPM	Ambulatory Blood Pressure Monitoring
AE	Adverse Event
ANCOVA	Analysis of Covariance
BP	Blood Pressure
CIP	Clinical Investigation Plan
DBP	Diastolic Blood Pressure
DSMB	Data Safety Monitoring Board
eGFR	estimated Glomerular Filtration Rate
ESRD	End Stage Renal Disease
MAE	Major Adverse Events
MCMC	Markov Chain Monte Carlo
OBP	Office Blood Pressure
SAP	Statistical Analysis Plan
SBP	Systolic Blood Pressure
SV2	Screening Visit 2

3. Introduction

This document outlines the detailed statistical methods to be implemented for the data collected within the scope of the Medtronic Vascular SPYRAL HTN-ON MED Study: Global Clinical Study of Renal Denervation with the Symplicity Spyral™ Multi-Electrode Renal Denervation System in Patients with Uncontrolled Hypertension on standard medical therapy. The purpose of this study is to test the hypothesis that renal denervation decreases blood pressure and is safe when studied in the presence of up to three standard antihypertensive medications. Specifically, the SAP has the following purpose: to prospectively outline the types of analyses and presentations of data that will form the basis for conclusions to be reached that will answer the trial objectives outlined in the protocol, and to explain in detail how the data will be handled and analyzed, adhering to commonly accepted standards and

practices of biostatistical analysis in the medical device industry. Results obtained from the analyses outlined in this document will be the basis of the Clinical Study Report for this trial.

4. Study Objectives

The objective of this study is to test the hypothesis that renal denervation is safe and reduces systolic blood pressure (SBP) in patients with uncontrolled hypertension on one, two, or three standard antihypertensive medications compared to a sham-controlled population. In this study, "uncontrolled hypertension" is defined as an office systolic blood pressure (SBP) ≥150 mmHg and <180 mmHg, an office Diastolic Blood Pressure (DBP) ≥90 mmHg and a 24-hour Ambulatory Blood Pressure Monitoring (ABPM) average SBP ≥140 mmHg to <170 mmHg, all of which are measured at screening visit 2. Data obtained will be used to confirm the effect of renal denervation on elevated blood pressure in patients on 1, 2 or 3 anti-hypertensive medications. Data collected during the SPYRAL HTN-ON MED trial may be used to gain market approval or additional indications for the Symplicity Spyral™ multi-electrode renal denervation catheter (Symplicity Spyral™ catheter) and the Symplicity G3™ renal denervation RF generator from regulatory entities, including, but not limited to: the Pharmaceuticals and Medical Device Agency (PMDA), Health Canada, and the U.S. Food and Drug Administration (FDA).

5. Study Endpoints

5.1. Primary Endpoints

There are two primary endpoints in this study (one efficacy and one safety). The study will be considered successful if both the primary safety and efficacy endpoint hypotheses are met.

Powered Primary Efficacy Endpoint:

Baseline adjusted change (using analysis of covariance, ANCOVA) in SBP from baseline (screening visit 2, SV2) to 6 months post-procedure as measured by 24-hour ABPM.

Powered Primary Safety Endpoint:

Incidence of Major Adverse Events (MAE), defined as a composite of the following events, through one-month post-randomization (6 months for new renal artery stenosis):

- All-cause mortality
- End-Stage Renal Disease (ESRD)
- Significant embolic event resulting in end-organ damage
- Renal artery perforation requiring intervention
- Renal artery dissection requiring intervention
- Vascular complications
- Hospitalization for hypertensive crisis not related to confirmed non-adherence with medications and/or the protocol
- New renal artery stenosis >70%, confirmed by angiography and as determined by the angiographic core laboratory

5.2. Secondary Endpoints

Secondary Efficacy Endpoints

- Change in SBP from baseline (screening visit 2) as measured by 24-hour ABPM at 3, 6, 12, 24 and 36 months post-procedure
- Change in office systolic blood pressure from baseline (screening visit 2) at 1, 3, 6, 12, 24 and 36 months post-procedure
- Change in diastolic blood pressure from baseline (screening visit 2) as measured by 24-hour ABPM at 3, 6, 12, 24 and 36 months post-procedure
- Change in office diastolic blood pressure from baseline (screening visit 2) at 1, 3, 6, 12, 24 and 36 months post-procedure
- Incidence of achieving target office systolic blood pressure (SBP <140 mmHg) at 1, 3, 6, 12, 24 and 36 months post-procedure

Secondary Safety Endpoints

Acute/procedural safety at 1-month post-procedure

- Significant embolic event resulting in end-organ damage
- Renal artery perforation requiring intervention
- Renal artery dissection requiring intervention
- Vascular complications
- End-stage renal disease
- ≥40% decline in eGFR
- Increase in serum creatinine >50% from screening visit 2
- New myocardial infarction
- New stroke
- Renal artery re-intervention
- Major bleeding according to TIMI definition (i.e. intracranial hemorrhage, ≥5g/dl decrease in hemoglobin concentration, a ≥15% absolute decrease in hematocrit, or death due to bleeding within 7 days of the procedure)
- New renal artery stenosis >70%, confirmed by angiography and as determined by the angiographic core laboratory
- Hospitalization for hypertensive crisis not related to confirmed non-adherence with medication and/or protocol

Chronic Safety Secondary Endpoints at 3, 6, 12, 24 and 36 months post-procedure

- Composite Safety Endpoint, defined as a composite of the following events:
    - All-cause mortality
    - End-stage renal disease (ESRD)
    - Significant embolic event resulting in end-organ damage
    - Renal artery perforation requiring intervention
    - Renal artery dissection requiring intervention
    - Vascular complications
    - Hospitalization for hypertensive crisis not related to confirmed non-adherence with medications and/or the protocol
    - New renal artery stenosis >70%, confirmed by angiography and as determined by the angiographic core laboratory

- ≥40% decline in eGFR
- Increase in serum creatinine >50% from screening visit 2
- New myocardial infarction
- New stroke
- Renal artery re-intervention
- Major bleeding according to TIMI definition (i.e. intracranial hemorrhage, ≥5g/dl decrease in hemoglobin concentration, a ≥15% absolute decrease in hematocrit, or death due to bleeding within 7 days of the procedure)
- Hospitalization for hypertensive crisis not related to confirmed non-adherence with medication and/or protocol

Summary of Health-related Quality of Life (HRQoL) analysis based on reporting measures using accepted QoL instruments (EQ5D)

Additional analyses

The following additional analyses will be conducted:

- Antihypertensive medication usage throughout the study, including escape patients and subjects with medication changes within 6-month follow-up.
- Additional procedural characteristics e.g. treatment duration, frequency of distal renal artery treatment, ablations per vessel, location of ablations, number of ablations per patient and other characteristics will be analyzed to assess their impact on blood pressure.
- Medication adherence will be assessed using results from drug testing. In addition, we will perform analyses to evaluate the effect of medications adherence on BP.
- Analyses looking at long term imaging will be performed.

6. Investigation Plan

The SPYRAL HTN-ON MED study is a multi-center, international, prospective, blinded, randomized, interventional, sham-controlled study. The goal is to demonstrate that catheter-based renal denervation using the Symplicity Spyral catheter and the Symplicity G3 generator is an effective and safe treatment for hypertension when studied in the presence of one, two, or three anti-hypertensive medication classes. Subjects must be on at least 50% of the maximum manufacturer's dosage of antihypertensive medication. Anti-hypertensive medication classes must include a thiazide-type diuretic, a dihydropyridine calcium channel blocker, an ACE-I/ARB, or a beta blocker, (when prescribed with other qualifying medications, 12.5 mg hydrochlorothiazide is acceptable as the minimum dosage).

7. Randomization and Blinding

Randomization will be stratified by study center at a 2:1 (RDN to control) ratio to:

- Denervation group (RDN): Subjects remain blinded and are treated with the renal denervation procedure.
- Control group: Subjects remain blinded and remain on the catheterization lab table for at least 20 minutes prior to introducer sheath removal.

Investigational sites will access randomization allocation via a password-protected system that can only be accessed by those approved by the study sponsor.

All study staff and necessary hospital personnel will be instructed that subjects are not to be informed of their randomization assignments and appropriate measures should be taken to minimize the risk of premature unblinding.

The Investigator performing the catheterization lab procedures and his/her designated study staff will be blinded to a subject's randomization group up until the angiography is completed and inclusion/exclusion confirmed following the angiography. However, investigators performing study follow-up visits and the subject's referring/managing physicians will not be proactively informed of a subject's treatment assignment to minimize potential bias in the subject's care decisions. To minimize potential bias in the measurement of Office BP and ABPM, each investigational site will specify several designated "blinded" members of their study staff that will not be informed of the subject's group assignments and will be responsible for performing the office blood pressure measurements, conducting ABPM preparation and printing results upon a patient completing the ABPM. Prior to unblinding, the effectiveness of blinding will be assessed by asking blinded study staff which group they believe the subject was randomized to.

Subjects will be blinded during the renal angiogram by a combination of conscious sedation, sensory isolation (e.g., blindfold and music), and lack of familiarity with the procedural details and duration (i.e., subjects will not know the difference between the renal angiography procedure alone and the renal angiography and denervation procedure). Subjects will continue to be blinded by only interacting with blinded site personnel through the 6-month follow-up visit post-procedure. Blinding effectiveness will be assessed by asking the subject which group they believe they were randomized to. All subjects will be unblinded after the completion of their required 6-Month follow-up testing.

8. Determination of Sample Size

This study will be conducted as an adaptive Bayesian trial with an informative prior. A Bayesian power prior approach [3,4] in conjunction with a discount function will be used to incorporate the prior data. The discount function reduces the strength of the prior data if disagreements are observed with the prospective data.

The prior data consists of the first consecutively randomized 80 subjects in the SPYRAL HTN-ON MED study, which were randomized in a 1:1 ratio to RDN or control. The results from these 80 subjects have already been analyzed and published [6].

The prospective data consists of the following two cohorts:

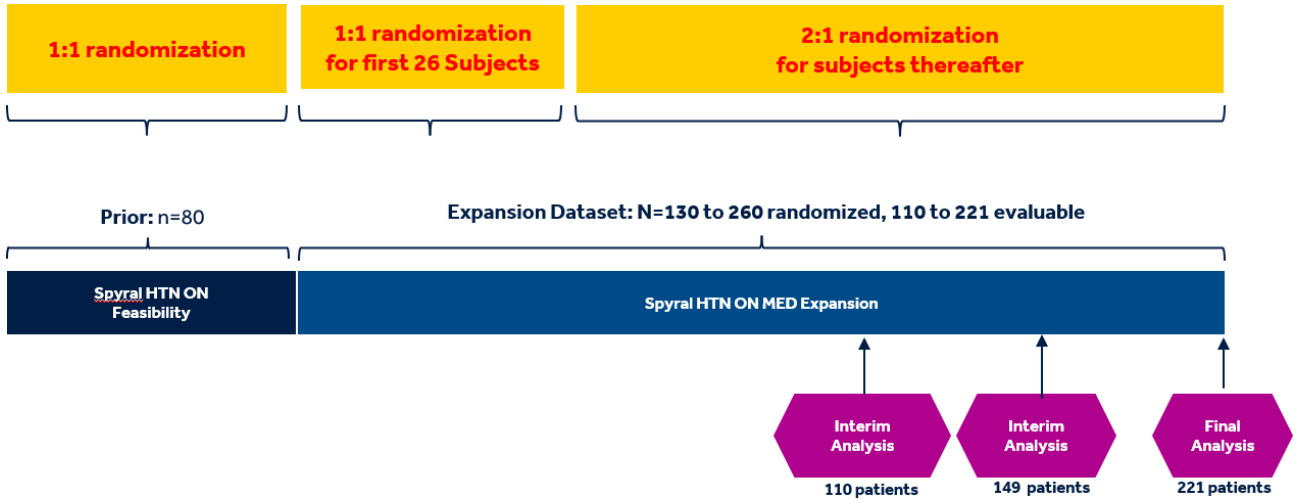
1. Subjects 81 to 106 which have already been enrolled and were randomized in a 1:1 ratio to RDN vs. control, and whose data has not been unblinded and analyzed by Medtronic.
2. All remaining subjects in the SPYRAL HTN-ON MED Study from 107 onwards which will be randomized in a 2:1 ratio to RDN vs control.

The weight of the prior data will be adjusted using a discount function, which scales from 0 to 1, according to the similarity of the prior and prospective data. The discount function adjusts the amount of weight the prior receives. This prevents the use of an informative prior where exchangeability issues are present (e.g., the prior and prospective data are quite different). This discount function approach was proposed by the Medical Device Innovative Consortium (MDIC) working group and is a collaborative

effort between FDA and industry through the MDIC [1,2]. If the analyses show a high level of agreement for prospective data compared to the prior, the prior will be weighted at or near 100%. If the prospective data perform much worse than or much better than the prior, then the prior will receive very little or zero weight. The Bayesian adaptive design is set up to enroll patients until a sufficient sample size is achieved to have high probability of meeting the endpoint.

The sample size of the study will vary from a minimum of 110 to 221 subjects with evaluable data due to the adaptive nature of the trial. This will require approximately 130 to 260 randomized subjects assuming a 15% rate of attrition at 6-months.

- The first interim analysis will take place when we have a minimum of 110 subjects with evaluable data. With an expected attrition rate of 15%, this will require approximately 130 randomized subjects. If the attrition rate is higher than 15%, then additional subjects will be randomized in order to reach a minimum of 110 evaluable subjects.
- The second interim analysis will take place when we have a minimum of 149 subjects with evaluable data. With an expected attrition rate of 15%, this will require 175 randomized subjects. If the attrition rate is higher than 15%, then additional subjects will be randomized in order to reach a minimum of 149 evaluable subjects.
- If the study does not stop at the first or second interim analyses, then the final look will take place when 260 subjects have been randomized. This will result in 221 evaluable subjects based on a 15% attrition rate.



Simulations of trial design operating characteristics performed to demonstrate control of type I error and power are presented in section 9.9.4.3 for the primary efficacy endpoint.

9. Statistical Methods

9.1. Study Subjects

9.1.1. Disposition of Subjects

A subject disposition table will be provided for each follow-up visit containing the following information:

- The number of subjects who died or withdrew prior to each follow-up
- The number of subjects eligible for each follow-up visit
- The number of subjects completing each follow-up visit within the protocol specified window
- The number of subjects completing each follow-up outside the protocol specified window
- The number of subjects who did not complete their follow-up

9.1.2. Clinical Investigation Plan (CIP) Deviations

A study deviation is an event where the investigator or investigational site personnel did not conduct the clinical study according to the Clinical Investigation Plan or Clinical Study Agreement. The investigator is not allowed to deviate from the above mentioned documents except with prior approval and under emergency circumstances. All deviations shall be documented and explained, regardless the reason for the deviation. Medtronic will assess the significance of all deviations and evaluate the need to amend the Clinical Investigation Plan or to early terminate the investigation, in accordance with Medtronic SOPs.

9.1.3. Analysis Sets

9.1.3.1. Intention-To-Treat (ITT) Population

All randomized subjects analyzed according to their randomized treatment. Subjects who meet the anti-hypertensive medication escape criteria (office SBP ≥180 mmHg or <115 mmHg associated with symptoms of hypotension or safety concern requiring medication changes) will be analyzed using Last Observation Carried Forward (LOCF) for their blood pressure measurements. Safety outcomes, and office and ambulatory blood pressure outcomes at each follow-up visit will be presented for this population.

9.1.3.2. Modified Intention-To-Treat (mITT) Population

All randomized subjects analyzed according to their randomized treatment. Subjects who meet the anti-hypertensive medication escape criteria (office SBP ≥180 mmHg or <115 mmHg associated with symptoms of hypotension or safety concern requiring medication changes) will be excluded from this population. Office and ambulatory blood pressure outcomes out to 6 months follow-up will be presented for this population.

9.1.3.3. Per Protocol Population

All randomized subjects, meeting the following criteria:

1. Subjects showing medication compliance in blood and/or urine (via drug testing data) at SV2, 3-months, and 6-months, compared with prescribed AH Meds at SV2.
2. Exclude subjects with protocol deviation code 101 (consent not obtained).
3. Exclude subjects who do not meet the following Inclusion/Exclusion criteria:
    - Inclusion: Individual has an office SBP ≥150 mmHg and <180 mmHg and an office DBP ≥90 mmHg measured at SV2, according to the guidelines in Appendix L7 of the study protocol.

- Inclusion: Individual has a 24-hour ABPM average SBP ≥140 and <170 mmHg measured at SV2, according to guidelines in Appendix L7 of the study protocol.
- Exclusion: Individual has undergone prior renal denervation.
- Exclusion: Individual has renal artery anatomy that is ineligible for treatment.

4. Exclude subjects meeting the anti-hypertensive medication escape criteria (office SBP ≥180 mmHg OR <115 mmHg associated with symptoms of hypotension or safety concern requiring medication changes).
5. Exclude subjects who did not receive the treatment they were randomized to.

Office and ambulatory blood pressure outcomes out to 6 months follow-up will be presented for this population.

9.1.3.4. Medication Adherence Population 1

All randomized subjects, meeting the following criteria:

1. Subjects showing medication compliance in blood and/or urine (via drug testing data) at SV2, 3-months, and 6-months, compared with prescribed AH Meds at SV2.
2. Exclude subjects meeting the anti-hypertensive medication escape criteria (office SBP ≥180 mmHg OR <115 mmHg associated with symptoms of hypotension or safety concern requiring medication changes).
3. Exclude subjects who did not receive the treatment they were randomized to.

Office and ambulatory blood pressure outcomes out to 6 months follow-up will be presented for this population.

9.1.3.5. Medication Adherence Population 2

All randomized subjects, meeting the following criteria:

1. Subjects showing medication compliance in blood and/or urine (via drug testing data) at 3-months and 6-months, compared with drug testing AH Meds at SV2.
2. Exclude subjects meeting the anti-hypertensive medication escape criteria (office SBP ≥180 mmHg OR <115 mmHg associated with symptoms of hypotension or safety concern requiring medication changes).
3. Exclude subjects who did not receive the treatment they were randomized to.

Office and ambulatory blood pressure outcomes out to 6 months follow-up will be presented for this population.

9.1.3.6. Medication Adherence Population 3

All randomized subjects, meeting the following criteria:

1. Subjects showing medication compliance in blood and/or urine (via drug testing data) at 3-months and 6-months, compared with drug testing AH Meds at SV2. If a subject is compliant at 3M, but not compliant at 6M, then their 3M BP will be used at 6M.
2. Exclude subjects meeting the anti-hypertensive medication escape criteria (office SBP ≥180 mmHg OR <115 mmHg associated with symptoms of hypotension or safety concern requiring medication changes).
3. Exclude subjects who did not receive the treatment they were randomized to.

Office and ambulatory blood pressure outcomes out to 6 months follow-up will be presented for this population.

9.1.3.7. As Treated Population

All randomized subjects, analyzed according to the actual treatment received. Subjects randomized to RDN who do not get treated will be analyzed in the control arm. Subjects who meet the anti-hypertensive medication escape criteria (SBP≥180 mmHg or safety reasons requiring medication changes) will be analyzed using Last Observation Carried Forward (LOCF) for their blood pressure measurements. Office and ambulatory blood pressure outcomes out to 6 months will be presented for this population.

9.1.4. Crossover Procedures

Control subjects may crossover to receive renal denervation therapy after completing their 6 month follow-up visit. For the subjects who have already completed their 6-month visit at the time crossover procedures are available per protocol, the decision to crossover must take place at their next in-person visit (6, 12, 24 and 36-month follow-up or Unscheduled visit). All subjects will have 30 days from that visit to undergo the crossover procedure. Subjects that are more than 30 days from 6, 12, 24 or 36-month or unscheduled visit, must complete a crossover baseline visit prior to having the crossover renal denervation procedure. To crossover, the required baseline data must be collected, and the subject cannot meet any of the anatomical or eGFR exclusion criteria. Crossover subjects will undergo follow up visits at 1, 3, 6, 12, 24, and 36 months post-procedure. Subjects who do not meet required eligibility for crossover on the day of the procedure will undergo follow-up visits according to their original follow-up schedule. Crossover procedures will be offered once and will not be available at a later time if it was declined by the control subject during the allowed 30-day window.

9.2. General Methodology

Descriptive statistics of continuous outcomes will be presented by treatment group and include sample size, mean, median, standard deviation, minimum and maximum. For categorical outcomes, the number and percentage of subjects in each category will be presented by treatment group. Statistical comparisons between treatment groups will be made using the independent samples t-test for continuous outcomes and Fisher's exact test for categorical outcomes. Paired t-tests will be used to compare changes in blood pressure from baseline to follow-up within each treatment group. All statistical analyses will be performed using SAS for Windows (version 9.2 or higher) or other widely accepted statistical or graphical software. Patient data listings and tabular and graphical presentations of results will be provided. Unless otherwise specified, a two-sided 0.05 level of significance will be used to declare treatment groups significantly different.

9.3. Poolability Analyses

9.3.1. Poolability of study centers

The following analyses will be performed to evaluate the poolability of data from different study centers. If the resulting tests are significant at the 0.15 level, further exploratory analysis will be attempted to identify covariates that may explain differences. Otherwise, the data will be considered to be poolable across study centers.

- A logistic regression will be conducted, with Major Adverse Events (MAE) as the dependent variable, and treatment, study center, treatment * study centers as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary safety endpoint is considered consistent among the sites.
- A linear regression will be conducted, with change in SBP from baseline to 6 months as measured by 24-hour ABPM as the dependent variable, and baseline systolic ABPM, treatment, study center, treatment * study center interaction term as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary effectiveness endpoint is considered consistent among the study centers.

9.3.2. Poolability of US and Canada

The following analyses will be performed to evaluate the poolability of data from US and Canadian sites. If the resulting tests are significant at the 0.15 level, further exploratory analysis will be attempted to identify covariates that may explain differences. Otherwise, the data will be considered to be poolable across regions.

- A logistic regression will be conducted, with MAE as the dependent variable, and treatment, US/Canada, treatment * US/Canada as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary safety endpoint is considered consistent between US and Canadian regions.
- A linear regression will be conducted, with change in SBP from baseline to 6 months as measured by 24-hour ABPM as the dependent variable, and baseline systolic ABPM, treatment, US/Canada, treatment * US/Canada interaction term as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary effectiveness endpoint is considered consistent between US and Canadian regions.

9.3.3. Poolability of North America (US and Canada)/Rest of World (ROW)

The following analyses will be performed to evaluate the poolability of data from North America (NA) and Rest of World (ROW) sites. If the resulting tests are significant at the 0.15 level, further exploratory analysis will be attempted to identify covariates that may explain differences. Otherwise, the data will be considered to be poolable across regions.

- A logistic regression will be conducted, with MAE as the dependent variable, and treatment, NA/ROW, treatment * NA/ROW as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary safety endpoint is considered consistent between NA and ROW subgroups.
- A linear regression will be conducted, with change in SBP from baseline to 6 months as measured by 24-hour ABPM as the dependent variable, and baseline systolic ABPM, treatment, NA/ROW, treatment * NA/ROW interaction term as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary effectiveness endpoint is considered consistent between NA and ROW subgroups.

9.4. Handling of Missing Data and Dropouts

Every effort will be made to minimize missing data for the primary efficacy endpoint. A secondary analysis will be performed where missing data is imputed using SAS PROC MI. Missing 6-month outcomes will be imputed using baseline SBP, 3-month SBP, treatment group, age, gender and BMI. One hundred

imputed datasets will be generated using a MCMC (Markov Chain Monte Carlo) algorithm, and a pooled estimate of the treatment effect will be generated using SAS PROC MIANALYZE.

9.5. Adjustments for Multiple Comparisons

The primary safety and efficacy endpoints are independently powered and no adjustments for multiple comparisons will be made.

9.6. Demographic and Other Baseline Characteristics

Baseline variables will be tabulated. Categorical variables, including binary variables, will be reported by giving the number and percentage of patients in each category. Continuous variables will be reported by presenting the number of values, mean, standard deviation, median, minimum, and maximum value for each. No imputation will be performed for missing data unless otherwise stated.

9.7. Treatment Characteristics

Renal denervation treatment measures such as number of ablation attempts and number of generator codes will be summarized separately for each kidney, and for combined kidneys. Anti-hypertensive medication use will also be summarized at baseline and at each follow-up.

9.8. Interim Analyses

Interim analyses will be conducted and reviewed by the DSMB, along with an independent organization that will be performing the Bayesian analyses. Medtronic personnel will not have access to any unblinded results prior to the primary endpoint analyses. The interim analyses will take place when a minimum of N=110 and N=149 subjects have evaluable 6-month efficacy data, with a maximum study size of N=260 randomized subjects if the study does not stop at an interim look. With an expected attriation rate of 15% at 6-months, this will require approximately N=130 and N=175 randomized subjects at the first and second interim analyses.

1. The first interim analysis takes place when a minimum of 110 subjects have evaluable 6-month efficacy data, requiring approximately 130 randomized subjects to account for 15% attrition. If the attrition rate is higher than 15%, then additional subjects will be randomized in order to reach a minimum of 110 evaluable. The Bayesian efficacy analysis will be performed and P[suc] will be calculated, where P[suc] is the probability of accepting the alternative efficacy endpoint hypothesis, $P(\mu < 0 | \boldsymbol{y}, \boldsymbol{y_0}, \hat{\theta}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k))$, and is defined in detail in section 9.9.3.
    a. If P[suc] > 0.975 then the study has met the efficacy hypothesis and enrollment will be stopped. Any additional subjects that have been enrolled before the decision is made to stop for efficacy will be pooled with the existing subjects and analyzed as a secondary cohort.
    b. We calculate the probability of futility based on the maximum study size of 260 randomized subjects which requires us to impute the outcomes for subjects who have not yet been enrolled (see sections 9.8.2 and 9.8.3 below for details). If the posterior probability of futility from this calculation is <0.05 for the primary efficacy endpoint, then the study will have met the futility boundary and enrollment will be stopped. Any

additional subjects that have been enrolled before the decision is made to stop for futility will be pooled with the existing subjects and analyzed as a secondary cohort.

c. If the stopping rules in a and b above are not met, then we continue enrolling subjects to the second interim analysis.

2. If we don't stop for efficacy or futility at the first interim analysis, then enrollment will continue until the second interim analysis when a minimum of 149 subjects have evaluable 6-month efficacy data, requiring 175 randomized subjects to account for attrition. If the attrition rate is higher than 15%, then additional subjects will be randomized in order to reach a minimum of 149 evaluable. The Bayesian efficacy analysis will be performed and P[suc] will be calculated.
   a. If P[suc] > 0.975 then the study has met the efficacy hypothesis and enrollment will be stopped. Any additional subjects that have been enrolled before the decision is made to stop for efficacy will be pooled with the existing subjects and analyzed as a secondary cohort.
   b. We calculate the probability of futility based on the maximum study size of 260 randomized subjects which requires us to impute the outcomes for subjects who have not yet been enrolled (see sections 9.8.2 and 9.8.3 below for details). If the posterior probability of futility from this calculation is <0.05 for the primary efficacy endpoint, then the study will have met the futility boundary and enrollment will be stopped. Any additional subjects that have been enrolled before the decision is made to stop for futility will be pooled with the existing subjects and analyzed as a secondary cohort.
   c. If the stopping rules in a and b above are not met, then we continue enrolling subjects to the final analysis.
3. If we don't stop for efficacy or futility at the second interim analysis, then enrollment will continue until the maximum study size of 260 randomized subjects. This will result in 221 evaluable subjects based on a 15% attrition rate. The Bayesian efficacy analysis will be performed and P[suc] will be calculated for the primary efficacy enpdoint.
   a. If P[suc] > 0.975 then we have met the primary efficacy endpoint hypothesis.

9.8.1. Mathematical forms for success:

$$P(\mu < 0 | \boldsymbol{y}, \boldsymbol{y_0}, \hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)) > 0.975$$

where y and y_0 represent the prospective data and prior data respectively, the notation $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ is used to denote that the estimate of the discounting parameter $\hat{\alpha}$, which depends on the prospective data, prior data, and the Weibull shape and scale parameters, and $\mu = \mu_t - \mu_c$ represents the baseline-adjusted treatment effect of BP change comparing RDN and control groups. See section 9.9.3 for more details.

9.8.2. Mathematical forms for futility:

$$\frac{\sum_1^{N_{rep}} I(P(\mu < 0 | y_{imp}) > 0.975)}{N_{rep}} < 0.05$$

Where y_{imp} is a single complete dataset after imputation, N_{rep} is the number of imputation simulations done and $\mu = \mu_t - \mu_c$ represents the baseline-adjusted treatment effect of BP change comparing RDN and control groups where μ_t and μ_c are the baseline adjusted BP changes in the RDN and control groups respectively.

9.8.3. Futility imputation procedure

For the futility calculation, we calculate the probability of futility based on the maximum study size of 260 randomized subjects which means we have three types of subjects to consider:

1. Subjects who have reached their 6-month endpoint.
2. Subjects who are enrolled and have baseline blood pressure but have not reached their 6 month endpoint.
3. Subjects who have not been enrolled.

Imputation procedures will be used to impute the blood pressure change for subjects of type 2 and type 3 above.

For type 2 subjects we use the following procedure:

2.1 Construct the posterior predictive distribution for blood pressure change using type 1 subjects, incorporating the prior data.
2.2 Simulate samples from the predictive distribution to impute the missing values of blood pressure change for type 2 subjects, conditional on observed baseline blood pressure.

For type 3 subjects, we do the following:

3.1 Construct the posterior predictive distribution for blood pressure change using type 1 subjects, including the prior data.
3.2 Simulate baseline blood pressure for type 3 patients.
3.3 Simulate blood pressure change from the predictive distribution.
3.4 Impute blood pressure change for type 3 subjects, conditional on simulated baseline blood pressure.

We then combine the type 1 subjects with the imputed subjects from steps 2.2 and 3.4 into a single dataset and construct the endpoint using the BayesDP function. This procedure will be repeated for many datasets ($N_{rep} = 1000$) and we calculate the number of times the alternative hypothesis is accepted. If the proportion of times the alternative hypothesis is accepted is less than 5%, stop enrollment due to futility.

We will use v1.3.2. of the bayesDP package, available on the Comprehensive R Archive Network (CRAN) [https://CRAN.R-project.org/].

9.9. Evaluation of Objectives

The study will be considered successful if we meet both the primary safety and efficacy hypotheses.

9.9.1. Primary Safety Endpoint

The primary safety endpoint of the study is the incidence of Major Adverse Events (MAE), defined as a composite of the following events through one month post-randomization (6 months for new renal artery stenosis):

- All-cause mortality
- End-stage renal disease (ESRD)
- Significant embolic event resulting in end-organ damage
- Renal artery perforation requiring intervention

- Renal artery dissection requiring intervention
- Vascular complications
- Hospitalization for hypertensive crisis not related to confirmed non-adherence with medications or the protocol
- New renal artery stenosis >70%

The primary safety analysis will be performed using the ITT population defined in 9.1.3.1.

9.9.1.1. Primary Safety Endpoint Analysis

Medtronic is using a performance goal approach to power the primary safety endpoint.

The safety performance goal for the Major Adverse Event (MAE) rate was developed based on review of and comparison to event rates of other renal interventions. The review of renal intervention literature reported event rates of between 3.6 and 17.2%. The reported events differed among the studies; however, for a subset of these studies, we could estimate rates for a composite of events similar to our protocol's MAE composite (see **Table 1** below). The major adverse event rate from these studies was 7.1%.

Table 1: MAE Rates of Literature Reported Studies

	MAE Rate
ROCHA[1]	4.8%
ASTRAL[2]	10.1%
Bax[3]	17.2%
Van Jaarsveld[4]	3.6%
Laird[5]	8.0%
Coral[6]	5.1%
Bradaric[7]	5.7%

[1] Rocha-Singh K, Jaff MR, Rosenfield K, ASPIRE-2 Trial Investigators. Evaluation of the safety and effectiveness of renal artery stenting after unsuccessful balloon angioplasty: the ASPIRE-2 study. J Am Coll Cardiol. 2005 Sep 6;46(5):776-83.

[2] Wheatley K, Ives N, Gray R, et al. Revascularization versus medical therapy for renal-artery stenosis. N Engl J Med 2009; 361:1953–1962.

[3] Bax L, Woittiez AJ, Kouwenberg HJ, et al. Stent placement in patients with atherosclerotic renal artery stenosis and impaired renal function: a randomized trial. Ann Intern Med 2009; 150:840–841.

[4] Van Jaarsveld, et al. The effect of balloon angioplasty on hypertension in atherosclerotic reanl-artery stenosis. N Engl J Med 2000; 342: 1007-14.

[5] Laird, et al. Safety and efficacy of renal artery stenting following suboptimal renal angioplasty for de novo and restenotic ostial lesions: results from a nonrandomized, prospective multicenter registry. J Vasc Interv Radiol 2010; 21: 627-637.

[6] Cooper CJ, Murphy TP, Cutlip DE, et al. Stenting and Medical Therapy for Atherosclerotic Renal-Artery Stenosis. NEJM 2014; 370(1): 13-22.

[7] Bradaric, C. et al. Drug-eluting stents versus bare-metal stents for the prevention of restenosis in patients with renovascular disease.

Jaff[8]	6.9%
Bersin[9]	9.8%
Overall Average (weighted by study size)	**7.1%**

The performance goal is set to be 7.1%, which is the meta-analysis rate from historical trials in **Table 1**.

The primary safety null and alternative hypotheses are:

H_0: $\pi \geq 7.1\%$ vs.

H_a: $\pi < 7.1\%$

where π is the MAE rate for patients undergoing renal denervation. Under the assumption that the true rate is 3.5%, and using a one-sided 0.05 level of significance, an evaluable sample size of 253 renal denervation patients yields 80% power to reject the null hypothesis in favor of the alternative. The exact binomial test was used for the sample size calculation for the primary safety endpoint hypothesis.

In other words, the primary safety endpoint hypothesis is designed to show whether the true MAE rate is lower than 7.1%. Compared to the literature reported event rates for renal intervention, we believe that these thresholds are appropriate for demonstrating safety of the device given the expected performance rates of similar renal intervention trials, particularly when balanced with the expected blood pressure reductions.

Medtronic proposes multiple sources of study patients as shown in **Table 2** below to ensure 253 patients treated with the Symplicity Spyral catheter (including branch treatment) are available for analysis. The first 253 subjects with evaluable safety data from the sources in Table 2 will be used to perform the primary safety endpoint analysis.

With a sample size of 253 and a one-sided significance level of 0.05, a maximum of 11 subjects with MAE will enable us to meet the safety primary endpoint, resulting in an event rate of 4.3% with a one-sided 95% upper confidence bound of 7.09% using the exact binomial method. The primary safety endpoint analysis will only be performed once using the first 253 subjects as detailed in this section. We will continue to report safety outcomes for all study subjects under secondary safety objectives as outlined in section 9.9.2.

Table 2: Study Sources of Patients for Primary Safety Endpoint Data

Study
SPYRAL HTN-OFF MED prior data (First 80 Subjects) Randomized 1:1 to RDN:CONTROL

[8] Jaff MR, Bates M, Sullivan T, et al. Significant reduction in systolic blood pressure following renal artery stenting in patients with uncontrolled hypertension: results from the HERCULES trial. Catheter Cardiovasc Interv. 2012 Sep 1;80(3):343-50.

[9] Bersin RM, Ansel G, Rizzo A, et al. Nine-Month Results of the REFORM Study: A Prospective, Single-Arm, Multicenter Clinical Study of the Safety and Effectiveness of the Formula Balloon-Expandable Stent for Treatment of Renal Artery Stenosis. Catheterization and Cardiovascular Interventions 82:266–273 (2013).

SPYRAL PIVOTAL – SPYRAL HTN-OFF MED Randomized 1:1 to RDN:CONTROL
SPYRAL HTN-ON MED First 106 Randomized 1:1 to RDN:CONTROL
SPYRAL HTN ON MED Extension Randomized 2:1 to RDN:CONTROL
SPYRAL HTN-OFF MED Crossovers (from prior and Pivotal)
SPYRAL HTN-ON MED Crossovers (from first 106 and Extension)

9.9.2. Secondary Safety Objectives

The following secondary safety endpoints will be assessed:

- Acute/Procedural Safety Secondary Endpoints – Compared Between Groups at 1 Month Post-Procedure
    - Significant embolic event resulting in end-organ damage
    - Renal artery perforation requiring intervention
    - Renal artery dissection requiring intervention
    - Vascular complications
    - End-stage renal disease (ESRD)
    - ≥40% decline in eGFR
    - Increase in serum creatinine >50% from screening visit 2
    - New myocardial infarction
    - New stroke
    - Renal artery re-intervention
    - Major bleeding according to TIMI definition (i.e. intracranial hemorrhage, ≥5g/dl decrease in hemoglobin concentration, a ≥15% absolute decrease in hematocrit, or death due to bleeding within 7 days of the procedure)
    - New renal artery stenosis >70%
    - Hospitalization for hypertensive crisis not related to confirmed non-adherence with medication and/or protocol
- Chronic Safety Secondary Endpoints - Compared Between Groups at 3, 6, 12, 24, and 36 months post-randomization.
    - Composite Safety Endpoint, defined as a composite of the following events:
        - All-cause mortality
        - End-stage renal disease (ESRD)
        - Significant embolic event resulting in end-organ damage
        - Renal artery perforation requiring intervention
        - Renal artery dissection requiring intervention
        - Vascular complications
        - Hospitalization for hypertensive crisis not related to confirmed non-adherence with medications and/or the protocol
        - New renal artery stenosis >70%

- ≥ 40% decline in eGFR
- Increase in serum creatinine >50% from screening visit 2
- New myocardial infarction
- New stroke
- Renal artery re-intervention
- Major bleeding according to TIMI definition (i.e. intracranial hemorrhage, ≥5g/dl decrease in hemoglobin concentration, a ≥15% absolute decrease in hematocrit, or death due to bleeding within 7 days of the procedure)
- Hospitalization for hypertensive crisis not related to confirmed non-adherence with medication and/or protocol
- Summary of Health-related Quality of Life (HRQoL) analysis based on reporting measures using accepted QoL instruments (EQ5D)

The safety endpoints will be adjudicated by the Clinical Events Committee (CEC). The following algorithm will be used to evaluate the safety event rates: The denominator will include all subjects who either had a CEC adjudicated event prior to the time of interest (180 days for 6 months events, for example), or had last contact date that is beyond the lower window of the follow up (166 days for 6 month events, for example). The numerator will include all subjects who had CEC adjudicated events up to the time of interest (180 days for 6 months events, for example).

The secondary safety endpoints, out to 6-months follow-up, will be compared between RDN and control groups using Fisher's exact test. Two-sided 95% confidence intervals of the difference between treatment groups will also be presented.

After 6-months follow-up, control subjects may crossover (undergo renal denervation), and secondary safety endpoints will be summarized by group (RDN, Crossovers, Non-Crossovers). RDN vs. Crossover vs. Non-Crossover groups will be compared out to 36 months post-procedure using chi-square tests for categorical data and ANOVA for continuous data.

The secondary safety analyses will be performed using the ITT population defined in section 9.1.3.1.

9.9.2.1. Renal Artery Stenosis Evaluation at 12 Months

With an expected rate of 3.1% for renal artery stenosis at 12 months [7], a sample size of 50 subjects will provide a 95% confidence interval of approximately (0.5%, 13.7%) using the exact method (calculated using an event rate of 2/50=4%).

Descriptive statistics of this endpoint at 12 months will be provided using counts, percentages and the 95% confidence interval.

9.9.3. Primary Efficacy Endpoint

The primary efficacy endpoint of the study is the baseline adjusted (ANCOVA) change in SBP from baseline (SV2) to 6-months post-procedure as measured by 24-hour ABPM.

In the context of an ANCOVA linear regression model, $\mu = \mu_t - \mu_c$ represents the baseline-adjusted treatment effect of BP change comparing RDN and control groups where μ_t and μ_c are the baseline adjusted BP changes in the RDN and control groups respectively. Let $\boldsymbol{y} = \{\boldsymbol{y_t}, \boldsymbol{y_c}\}$ and $\boldsymbol{y}_0 = \{\boldsymbol{y_{0t}}, \boldsymbol{y_{0c}}\}$ represent the prospective data and prior data respectively, where t = RDN group and c = control group. Let the hypotheses for the study be the following:

$$H_0: \mu = 0$$

$$H_a: \mu < 0$$

We reject H_0 if the probability is greater than 97.5%, i.e.

$$P(\mu < 0 | \boldsymbol{y}, \boldsymbol{y_0}, \hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)) > 0.975$$

where the notation $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ is used to denote that the estimate of $\hat{\alpha}$ depends on the prospective data, prior data, and the Weibull shape and scale parameters. In conjunction with a pre-specified decision rule controlling the prior data weight, the estimate of $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ represents a measure of similarity between prospective and prior data. Alternatively, in the absence of $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$, i.e., $P(\mu < 0 | \boldsymbol{y}, \boldsymbol{y_0})$, full weight would be given to the prior data.

9.9.4. Primary Efficacy Endpoint Analysis

The primary analysis of the powered primary endpoint will be performed using Bayesian methods as outlined in this section. The Bayesian posterior treatment effects will be determined along with the 95% Bayesian Credible Interval (BCI). The ITT population will be used as the primary analysis population for this endpoint. Secondary Bayesian effectiveness analyses will also be performed using the modified ITT, per-protocol and as-treated populations defined in section 9.1.3.

The power prior discount function approach is used to estimate μ, and determine $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$, the strength of the prior data used to estimate μ. $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ ranges from 0 to 1, where 1 means that 100% of the prior data is used and 0 means that no prior data is used. Before beginning the study, an initial value is chosen for $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$, call this value α_{max}. This α_{max} value is the maximum strength the prior data can receive. We intend to use the same enrollment criteria for the prior and prospective studies, and therefore believe that a value of $\alpha_{max} = 1$ is appropriate.

At interim looks and at the final analysis, we analyze the data using the power prior discount function method, this method will discount α_{max} to an appropriate value $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ where $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y}_0, \lambda, k) \leq \alpha_{max}$. This discounting is based on the discount function which is discussed in detail in the next section.

Under the adaptive procedure, if the prospective data diverges from the prior data at an interim look, the discount function will discount the strength of the prior data, thus requiring continued enrollment to maintain power to achieve the endpoint. Alternatively, if the prior and prospective data agree, there will be a smaller penalty from the discount function, thus fewer prospective patients would be needed to maintain power.

9.9.4.1. Discount Function Estimation Method

The power prior discount function method is comprised of four steps: **Compare, Discount, Combine, and Estimate.**

Compare:

We start by stochastically comparing prospective data vs prior data as follows,

For each treatment group, we separately fit the model to the combined prior and pivotal data:

$$y_i = \beta_0 + \beta_1 I(i \in \text{prior}) + \beta x_{2\ i} + \varepsilon_i, \quad \varepsilon_i \sim N(0, \tau^2),$$

where $I(i \in \text{prior}) = 1$ if the subject is from the prior dataset, and 0 otherwise, y_i is the BP change for the ith observation and x_i is the mean centered baseline BP for the ith observation. With flat priors on

each parameter, we estimate the posterior probability that $\beta_1 > 0$ by first computing, using Monte Carlo sampling

$$p^* = P[\tilde{\beta}_1 > 0 \mid \boldsymbol{y}, \boldsymbol{y_0}].$$

Having calculated this separately for both the RDN (p_t^*) and control groups (p_c^*), they are transformed to p_t and p_c using

$$p = \begin{cases} 2p^*, & p^* \leq 0.5 \\ 2(1-p^*), & p^* > 0.5 \end{cases}$$

Now, under this transformation, if p_t or p_c are close to 0, there is a high probability that the prospective data and prior data come from different populations and discounting should be applied to reduce the influence of the prior. On the other hand, if p_t or p_c are close to 1, there is a high probability that the prospective data and prior data come from similar populations and minimal discounting should be applied.

Discount:

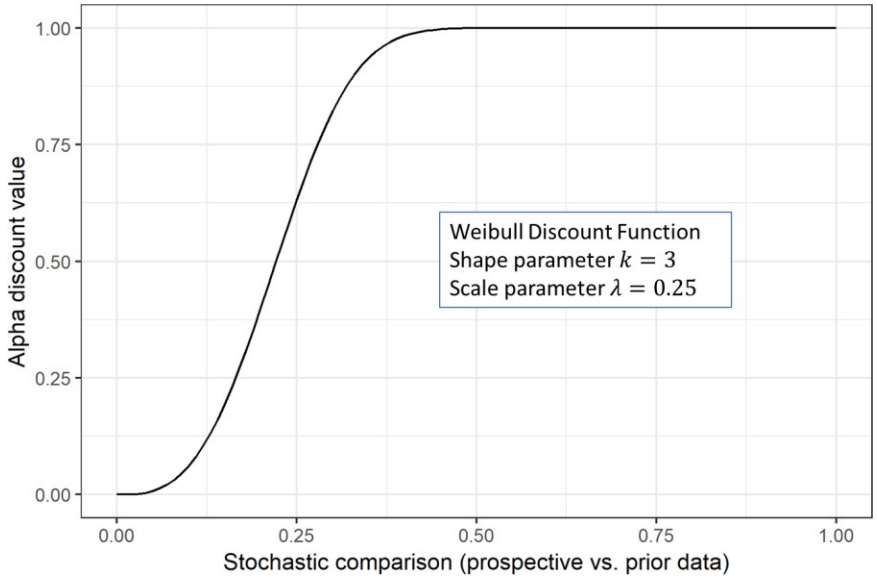
We discount α_{max} based on the value of p_t and p_c from the **Compare** step and the discount function $F(p)$,

$$\hat{\alpha} = \alpha_{max} F(p),$$

where $F(p)$ is a function between 0 and 1. A two-sided Weibull function will be utilized as follows:

$$F(p) = 1 - e^{-\left(\frac{p}{\lambda}\right)^k}$$

For this study, we will be using a shape parameter of $k = 3$ and a scale parameter of $\lambda = 0.25$ (illustrated below). Note that we will use the same Weibull function parameters for both RDN group and control group, but p_t and p_c will have different values from the **Compare** step.



Combine:

Using the power prior method and $\hat{\alpha}$ we can combine the prior and prospective data together using Bayesian techniques to construct the posterior distribution for μ as follows. We first begin with a hierarchical linear regression model of the form:

$$\begin{aligned} y_i \mid \mu_t, \mu_c, \beta, \sigma^2 &\sim N(\mu_t I(i \in t) + \mu_c I(i \in c) + x_i \beta, \sigma^2), i = 1, \ldots, n, \\ \mu_t \mid \mu_{0t}, \sigma^2_{0t}, \hat{\alpha}_t &\sim N(\mu_{0t}, \sigma^2_{0t}/\hat{\alpha}_t), \\ \mu_c \mid \mu_{0c}, \sigma^2_{0c}, \hat{\alpha}_c &\sim N(\mu_{0c}, \sigma^2_{0c}/\hat{\alpha}_c), \\ \beta &\sim N(a_\beta, b^2_\beta), \\ \pi(\sigma^2) &\propto \sigma^{-2}, \end{aligned}$$

where $N(\cdot)$ denotes a normal distribution, $I(\cdot)$ is the indicator function where $I(i \in t)$ indicates that observation i is in the RDN group and $I(i \in c)$ indicates observation i is in the control group, y_i is the BP change for the ith observation, x_i is the mean centered baseline BP value for the ith observation, and each of μ_{0t}, μ_{0c}, σ^2_{0t}, σ^2_{0c} are hyperparameters estimated from the historical data, and $a_\beta = 0$ and $b_\beta = 10^{10}$.[10]

To carry out estimation in a computationally efficient manner, we rely on a reparameterization derived in Gelman [4]. First, construct vectors and matrices: $\boldsymbol{y} = (y_1, \ldots, y_n)^T$, $\boldsymbol{x}_t$ and $\boldsymbol{x}_c$ vectors of binary treatment control indicators taking values 0 and 1, $\boldsymbol{x}_\beta$ the vector of mean centered baseline values, $\boldsymbol{X} = (\boldsymbol{x}_t \mid \boldsymbol{x}_c \mid \boldsymbol{x}_\beta)$ the $n \times 3$ design matrix, and $\boldsymbol{\theta} = (\mu_t, \mu_c, \beta)^T$ parameters of interest to be estimated. Then, we can write the model as

$$\boldsymbol{y}_* \sim N(\boldsymbol{X}_* \boldsymbol{\theta}, \boldsymbol{\Sigma}_*),$$

where $\boldsymbol{y}_* = (\boldsymbol{y}^T, \mu_{0t}, \mu_{0c}, a_\beta)^T$, $\boldsymbol{X}_* = (\boldsymbol{X}^T, \boldsymbol{I}_3)^T$, $\boldsymbol{I}_3$ is a 3×3 identity matrix, and

$$\boldsymbol{\Sigma}_* = \begin{pmatrix} \sigma^2 \boldsymbol{I}_n & 0 \\ 0 & \boldsymbol{\Sigma}_\theta \end{pmatrix},$$

where

$$\boldsymbol{\Sigma}_\theta = \begin{pmatrix} \sigma^2_{0t}/\hat{\alpha}_t & 0 & 0 \\ 0 & \sigma^2_{0c}/\hat{\alpha}_c & 0 \\ 0 & 0 & b^2_\beta \end{pmatrix}.$$

The posterior mean of $\boldsymbol{\theta}$ is found via least squares as

$$\hat{\boldsymbol{\theta}} = (\boldsymbol{X}_*^T \boldsymbol{\Sigma}_*^{-1} \boldsymbol{X}_*)^{-1} \boldsymbol{X}_*^T \boldsymbol{\Sigma}_*^{-1} \boldsymbol{y}_*,$$

and the posterior variance of $\boldsymbol{\theta}$ is

$$\boldsymbol{V}_\theta = (\boldsymbol{X}_*^T \boldsymbol{\Sigma}_*^{-1} \boldsymbol{X}_*)^{-1}.$$

Thus, the posterior distribution of $\boldsymbol{\theta}$ is $\boldsymbol{\theta} \mid \boldsymbol{y}_*, \boldsymbol{\Sigma}_* \sim N(\hat{\boldsymbol{\theta}}, \boldsymbol{V}_\theta)$. Both $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{V}_\theta$ are composed of an unknown σ^2. The marginal posterior distribution of σ^2 is

[10] In the R package bayesDP version 1.3.2, the bdplm function parameter 'prior_covariate_sd' is described as "The prior standard deviation(s) of the covariate effect(s). Default value is 1e4." Upon inspection of the code, it appears the standard deviation is internally scaled by 1e6, thus yielding a more diffuse prior for β. Therefore, the actual value of b_β used in the analysis is 1e10.

$$q(\sigma^2 \mid \boldsymbol{y}) \propto \pi(\sigma^2)|\sigma^2 \boldsymbol{I}_n|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\boldsymbol{y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\theta}})^T \boldsymbol{\Sigma}_*^{-1} (\boldsymbol{y}_* - \boldsymbol{X}_*\hat{\boldsymbol{\theta}})\},$$

which does not have a known distributional form. Here, $\pi(\sigma^2)$ is the prior distribution of σ^2. Thus, to draw samples from the posterior distribution of σ^2, we rely on a grid search to sample values of σ^2.

We proceed by drawing samples from $\sigma^2 \sim q(\sigma^2|\boldsymbol{y})$. Using these estimates, we input them into the posterior distribution of $\boldsymbol{\theta}$. By repeating this process, a large number of times, we will be drawing posterior samples from $\boldsymbol{\theta}$ which will account for the uncertainty in both σ^2 as well as appropriately weighting the prior data based on $\hat{\alpha}_t$ and $\hat{\alpha}_c$.

We can than construct the following contrast of interest from the drawn samples:

$$\mu = \mu_t - \mu_c.$$

Use of this contrast leads to the univariate distribution of interest concerning the mean BP change difference between the RDN and control groups.

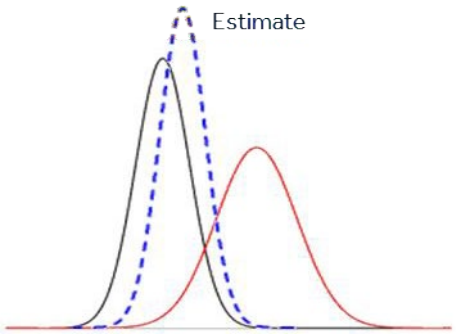
Estimate:

The posterior distribution from the combined prior and prospective data is used to estimate the posterior probability

$$P(\mu < 0 | \boldsymbol{y}, \boldsymbol{y_0}, \hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)) \qquad (1)$$

where the notation $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ is used to denote that the estimate of $\hat{\alpha}$ depends on the prospective data, prior data, and the Weibull shape and scale parameters. In conjunction with a pre-specified decision rule controlling the prior data weight, the estimate of $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$ represents a measure of similarity between prospective and prior data. Alternatively, in the absence of $\hat{\alpha}(\boldsymbol{y}, \boldsymbol{y_0}, \lambda, k)$, i.e., $P(\mu < 0 | \boldsymbol{y}, \boldsymbol{y_0})$, full weight would be given to the prior data.

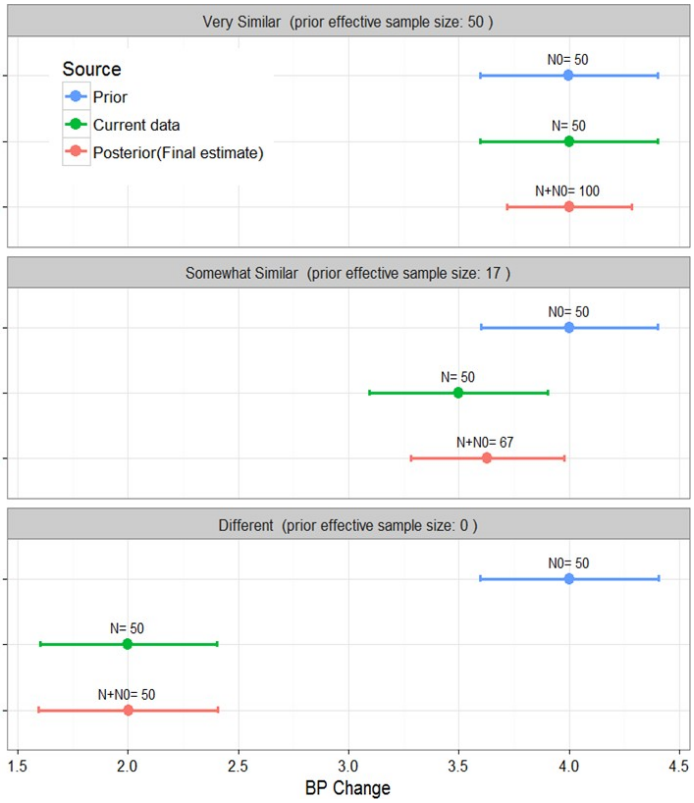
The blue dashed line in the figure below is an illustrative example of the estimate from prospective data (black) and prior data (red).



The analysis in (1) is performed at all interim looks and at the final analysis.

9.9.4.2. Illustration of Discount Function Scenarios

The Figure below shows how the discount function operates with hypothetical data sets



The panels in this figure can be interpreted as follows:

- Top panel: The prospective (current) data is very similar to the prior. The discount function allows for full strength of the prior. The posterior (final estimate) is a balance between the prior and prospective study.
- Middle panel: The prospective (current) data is similar to the prior. The discount function penalty is moderate, resulting in a prior effective sample size of 17 out of a max of 50. Because the agreement is reasonable, the posterior (final estimate) is similar to both the prior and prospective study.
- Bottom panel: The prospective (current) data shows lower performance than the prior. The discount function produces a substantial penalty resulting in no weight to the prior. The posterior (final estimate) is essentially the same as the prospective (current) estimate.

9.9.4.3. Simulation of Primary Efficacy Endpoint Operating Characteristics

Simulations were performed to assess operating characteristics for the primary efficacy endpoint and are presented in the two tables below. We used 8000 trial simulations to estimate the power and 15000 simulations to estimate the type I error. The overall power for the primary efficacy endpoint from **Table 4** is 96%, with a one-sided type I error rate of 0.038.

Table 3: Simulation Parameters for Primary Efficacy Endpoint

Prior Baseline Adjusted RDN Group Mean (SE)	-8.85 (1.75) mmHg
Prior RDN Group N	36
Prior Baseline Adjusted Control Group Mean (SE)	-1.80 (1.75) mmHg
Prior Control Group N	36

Maximum Prior Patients	36 + 36 = 72 (out of 80)
Prospective Study Expected Treatment Difference	-5.0 mmHg
Prospective Study RDN Group Mean (SD)	-6.8 (12) mmHg
Prospective Study Control Group Mean (SD)	-1.8 (12) mmHg
Weibull Discount Function Parameters	Shape: $k = 3$, Scale: $\lambda = 0.25$
Attrition rate at 6-months	15%

Table 4: Study Operating Characteristics for Primary Efficacy Endpoint

Trial Success Rate (Power)	96%
Type I Error	0.038
First Interim Look N	N=110 evaluable (~130 randomized)
Power at First Interim Look	83%
Second Interim Look N	N=149 evaluable (~175 randomized)
Power at Second Interim Look	91%
Maximum Study Size	N=221 evaluable (260 randomized)
% of Simulations that Stop for Futility	1.71%

9.9.5. Primary Efficacy Endpoint Sensitivity Analyses

The following sensitivity analyses will be performed for the primary efficacy endpoint to evaluate the consistence of the results. For each of these analyses we will report the posterior distribution mean of μ, the 95% BCI of μ, and the posterior probability of $\mu < 0$.

9.9.5.1. Bayesian ANCOVA Model with Single Treatment Effect Parameter using Discount Prior Approach

Using the classical ANCOVA parameterization

$$y_i = \beta_0 + \mu\, I(i \in t) + \beta_x x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2),$$

there is a single treatment effect, μ. Similar to the power prior methodology, we can determine an informative prior distribution for μ using the feasibility data. The model above is fitted to the feasibility data assuming flat prior distributions, and the marginal posterior distribution of μ extracted. A stochastic comparison is performed in the same manner as Section 9.9.4.1, whereby we pool the feasibility and expansion data (this time pooling over treatment arms as well) and fitting the model

$$y_i = \beta_0 + \beta_1 I(i \in \text{prior}) + \beta_2 x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

The same Weibull discounting function is applied as per the primary analysis (Section 9.9.4.1) to yield a discounting parameter. The Bayesian ANCOVA model is fitted to the expansion data using flat prior distributions for all parameters except μ, which has a distribution of

$$\mu \mid Y_0 \sim N(\hat{\mu}, \hat{\sigma}^2/\alpha)$$

Where $\hat{\mu}$ and $\hat{\sigma}^2$ are the posterior mean and standard error estimates from fitting the ANCOVA model to the feasibility data, and α is the discount parameter. Note that this model is invariant to mean centering of the baseline BP, and therefore we do not apply it.

9.9.5.2. Bayesian ANCOVA Model with Single Treatment Effect Parameter using Full Coefficient Vector Prior Distribution

This analysis uses the same ANCOVA model as 9.9.5.1. with separate priors applied to $\boldsymbol{\beta} = (\beta_0, \mu, \beta_x)'$.. The power parameter is then used to down-weight the influence of $\hat{\boldsymbol{\beta}}_{\mathbf{0}}$, the prior estimate. Note that we consider σ^2 to be a nuisance parameter, and therefore a non informative prior distribution is used in the pivotal data update step. This approach allows for the full covariance matrix of β to be downweighted. In the case of the linear regression normal model $Y = X'\beta + \epsilon, \epsilon \sim N(0, \sigma^2)$ with known σ^2 (estimated using the feasibility data), with $D_0 = \{(x_{0i}, y_{0i}); i = 1,2, \ldots n_0\}$ and starting with a uniform initial prior $\pi_0(\beta) \propto 1$ the power-prior for β simply reduces to

$$\pi(\beta|D_0, \alpha) = N\left((X_0'X_0)^{-1}X_0'Y_0, \frac{(X_0'X_0)^{-1}\sigma^2}{\alpha}\right)$$

9.9.5.3. Bayesian ANCOVA Model with Single Treatment Effect Parameter using Propensity Score Overlap

This model will use a propensity score overlap for discounting the overall treatment effect. Similar to the approaches proposed by (Wang et al. 2019, 2020; Chen et al. 2020) [9]-[11], here we explore the propensity score overlap between the feasibility and expansion cohorts. We will fit a propensity score model including baseline covariates for age, gender, body mass index, obstructive sleep apnea, history of type II diabetes, baseline eGFR, baseline 24-hour SBP, baseline 24-hour DBP, medication-compliance at baseline (determined via drug testing), and number of AH medications at baseline. We calculate the proportion overlap of the propensity score kernel density distributions between the feasibility and expansion cohorts (using the Overlapping R package: Pastore M et al [13]). We then run the Bayesian model in 9.9.5.1 using the overlap proportion (between 0 and 1, with 1 = complete overlap and 0 = no overlap) as the discount parameter.

9.9.5.4. Commensurate Prior Model with Beta Priors on the Alpha-Discount Parameters without Mean Centering of Baseline Covariates

The commensurate-power-prior approach (Hobbs et al., 2011 [12]) will be used instead of using the discount-function approach. The following prior and likelihood specifications will be used:

Using a vector notation for the classical ANCOVA parameterization in section 9.9.5.1,

- $Y_0 \sim N(\boldsymbol{X\beta_0}, \sigma_0^2)$, $Y \sim N(\boldsymbol{X\beta}, \sigma^2)$, where, $\boldsymbol{\beta} = (\beta_0, \mu, \beta_x)'$ and similarly for $\boldsymbol{\beta_0}$ and **X** is the design matrix. For simplicity, assume σ_0^2 known - take estimate s_0^2 from D_0.

- Priors:

$$\begin{aligned} \pi(\beta_0) &\propto 1 \\ \pi(\beta_x) &= Normal\,(\beta_0, \frac{1}{\tau}) \\ \pi(\sigma^2) &\propto \frac{1}{\sigma^2} \\ \pi(\alpha) &= Beta(\max(\log\tau, 1), 1) \\ \pi(\log\tau) &= Cauchy(0,30). \end{aligned}$$

- τ is the commensurability measure parameter for which a Cauchy(0,30) prior is assumed on the log-scale. Thus, the prior on τ is concentrated very close to 0. For $\log\tau \leq 1, \alpha \sim Uniform(0,1)$ and as τ decreases the prior for β gets flatter. While, as τ increase, the distribution of α tends to get heavier towards 1.
- Mean centering will not be applied to the baseline covariates

9.9.5.5. Discount Power Prior Extension of the t-test

Taking the outcome as BP change from baseline, we can apply a t-test to determine the treatment effect between arms. Generalizing the Bayesian disount power approach to this model is described in section 2.5 of Haddad et al [2]. This model does not adjust for baseline BP, so the issue of mean centering is redundant. This analysis will be performed using the 'bdpnormal()' function, in version 1.3.2. of the bayesDP package, available on the Comprehensive R Archive Network (CRAN) [https://CRAN.R-project.org/].

9.9.6. Secondary Analysis of Primary Efficacy Endpoint

As a secondary analysis of the powered primary efficacy endpoints, frequentist ANCOVA methods will be used to determine the baseline adjusted treatement effect for the study populations defined in section 9.1.3.

As outlined in section 9.8, the first and second interim analyses takes place when a minimum of 110 and 149 subjects have 6-month follow-up ABPM data available respectively. At each interim analysis, the Bayesian efficacy analysis will be performed and, if the pre-specified stopping rules are met, then enrollment will be stopped. If the study stops for efficacy or futility at either the first or second interim analysis, then any additional subjects that have been enrolled before the decision to stop has been made will not be part of the primary endpoint analysis, but instead will be pooled with the existing subjects and analyzed as a secondary cohort. This analysis will be performed using frequentist ANCOVA methods to determine the baseline adjusted treatment effect and will be performed by Medtronic, as Medtronic will be unblinded at this stage.

9.9.7. Secondary Efficacy Objectives

The following additional secondary efficacy endpoints will be assessed:

- Change in SBP from baseline (SV2) as measured by 24-hour ABPM at 3, 6, 12, 24 and 36 months post-procedure.
- Change in office SBP from baseline (SV2) at 1, 3, 6, 12, 24 and 36 months post-procedure.
- Change in DBP from baseline (SV2) as measured by 24-hour ABPM at 3, 6, 12, 24 and 36 months post-procedure.
- Change in office DBP from baseline (SV2) at 1, 3, 6, 12, 24 and 36 months post-procedure.

- Incidence of achieving target office SBP (SBP <140 mmHg) at 1, 3, 6, 12, 24 and 36 months post-procedure.

RDN vs. control groups will be compared out to 6-months post randomization, prior to the crossover procedure. Statistical comparisons will be performed using the independent samples t-test for continuous endpoints and Fisher's exact test for categorical endpoints. In addition, two-sided 95% confidence intervals of the difference between RDN and control groups will be presented. Changes in blood pressure measurements from baseline to follow-up within each treatment group will be assessed using paired t-tests. Two-sided 95% confidence intervals of the mean change from baseline will be presented for each treatment group. analysis of covariance (ANCOVA) models, adjusting the treatment effect for the baseline BP measurements will also be applied to all continuous secondary endpoints.

RDN vs. crossover vs. non-crossover groups will be compared out to 36 months post-procedure using chi-square tests for categorical data and ANOVA for continuous data.

Changes in blood pressure measurements from baseline to follow-up within each group will be assessed using paired *t*-tests. Two-sided 95% confidence intervals of the mean change from baseline will also be presented for each group.

The secondary efficacy analyses will be presented for all the study populations defined in section 9.1.3.

9.9.8. Additional Objectives

- Quality of Life (QOL) EQ-5D measures.
- Antihypertensive medication usage throughout the study, including escape patients and subjects with medication changes within 6-month follow-up.
- Additional procedural characteristics e.g. treatment duration, frequency of distal renal artery treatment, ablations per vessel, location of ablations, number of ablations per patient and other characteristics will be analyzed to assess their impact on blood pressure.
- Medication adherence will be assessed using results from drug testing, witnessed pill intake and BP protocol deviations. In addition, we will perform analyses to evaluate the effect of medication adherence on blood pressure change.
- Analyses looking at long term imaging will be performed.

RDN vs. control groups will be compared out to 6-months post randomization, prior to the crossover procedure. Statistical comparisons will be performed using the independent samples t-test for continuous endpoints and Fisher's exact test for categorical endpoints. In addition, two-sided 95% confidence intervals of the difference between RDN and control groups will be presented.

RDN vs. crossover vs. non-crossover groups will be compared out to 36 months post-procedure using chi-square tests for categorical data and ANOVA for continuous data.

Changes in continuous measurements from baseline to follow-up within each group will be assessed using paired t-tests. Two-sided 95% confidence intervals of the mean change from baseline will also be presented for each group.

The additional objectives will be analyzed using the ITT study population defined in section 9.1.3.1.

9.9.9. Win Ratio Analyses

Win Ratio analyses, which use hierarchical composite outcomes to compare the RDN and Control groups [14], will be used to analyze the blood pressure and AH medication data at follow-up. A prospective

secondary efficacy analysis will be performed at 6-months using the win ratio approach. The following endpoints will be included in the hierarchical comparison:

1. 24Hr SBP change from SV2 to 6M using a threshold of 5 mmHg
2. Medication burden (INDEX2) change from SV2 to 6M using a threshold of zero

The win ratio statistic is a generalisation of the non-parametric Wilcoxon rank-sum test for comparing continuous outcomes between 2 independent groups to multiple prioritized endpoints. Every RDN subject is paired with every control subject and analysed using the endpoints specified above.

a) For every pair:
   i. Calculate the change in the first endpoint (24Hr SBP) from SV2 to 6-months for the RDN subject (Δ_{RDN}) and the control subject (Δ_{CON})
   ii. Calculate the pairwise treatment effect for the pair: $\Delta_P=\Delta_{RDN}-\Delta_{CON}$
   iii. Compare Δ_P to the specified threshold (5 mmHg)
   iv. if the RDN subject has a better outcome compared to control subject using this threshold ($\Delta_P \leq -5$) then this results in a "win" for the RDN subject. We stop analyzing this pair and move on to the next
   v. if the control subject has a better outcome compared to the RDN subject ($\Delta_P \geq +5$) then this results in a "loss" for the RDN subject. We stop analyzing this pair and move on to the next
   vi. if the pairwise treatment effect is smaller than the threshold ($-5 < \Delta_P < +5$) or if either subject has missing data then this pair results in a "tie"
b) Only pairs classified as ties for the first endpint (24Hr SBP) proceed to the second hierarchical endpoint of medication burden change and we repeat step a) above using their medication burden change data and a threshold of zero.

After every pair has been analyzed, the win ratio statistic is calculated as the total number of wins divided by the total number of losses from both endpoints. The win ratio will be presented with a 95% confidence interval and p-value.

The ITT population, including the full feasibility cohort, will be used. For medication INDEX2, drug testing data will be used at SV2 and 6-months if available, otherwise the medication burden will be determined using prescribed medication data.

Two additional win ratio sensitivity analyses will be performed using thresholds of 3.5 and 0 mmHg for the 24Hr SBP endpoint and keeping the threshold of zero for the medication burden endpoint. The R package BuyseTest version 2.2.6 will be used to perform these analyses.

9.9.10. % Time in Target Range (TTR%) Analyses

Time in target range (TTR%) analyses [16], which calculate the percentage of time that a subject's BP is within pre-specified BP ranges will be calculated by performing linear interpolation over the days between successive BP measurements, including data from both the baseline and follow-up periods, and summarizing over the entire follow-up time period. Subjects must have a minimum of 2 BP measurements within the TTR time interval to be included in the calulation. All follow-up visits will be included in the TTR analyses. Office SBP TTR will be evaluated using a range of ≤140 mmHg and 24Hr SBP TTR will be evaluated using ranges of ≤140 & ≤130 mmHg. In addition, we will combine Office TTR (≤140 mmHg) with 24Hr SBP TTR (≤130 mmHg) by taking the maximum value for each subject within the TTR time interval.

For each subject we will calculate their TTR%, using the BP endpoints and BP ranges specified above, in the following time intervals:

- TTR% from baseline to 6-months
- TTR% from baseline to 12-months
- TTR% from baseline to 24-months

- TTR% from baseline to 36-months

These results will be tabulated and compared between treatment arms using non-parametric Kruskal-Wallis tests.

9.10. Safety Evaluation

Adverse Event (AE) information will be collected by the site from subject enrollment (consent) through study termination. AEs will be followed until the event has resolved (in the case of permanent impairment, the event will be followed until it stabilizes, and the overall clinical outcome has been ascertained).

The Investigator will report any adverse events that may occur to the Sponsor, and will assess seriousness, relationship (to the device, procedure and renal denervation therapy where applicable), subsequent intervention required, resolution status and whether or not the adverse event resulted in the subject's discontinuation from the study. The Investigator will provide further information regarding adverse events as requested by the Sponsor.

9.11. Subgroup Analyses

Analysis will be carried out for the following subgroups to assess consistency of results.

- Female vs. male gender
- Age at baseline <65 vs. ≥ 65 years
- BMI by tertiles (kg/m^2)
- Type 2 diabetics vs. non-diabetics
- Current smokers vs. former smokers vs non-smokers
- Baseline eGFR <60 vs. ≥60 mL/min/1.73 m^2
- Obstructive sleep apnea yes vs. no
- US vs. OUS subjects
- US African American vs. US non-African American subjects
- OUS European vs. Japanese vs. Australian subjects
- Baseline ambulatory SBP by tertiles and medians (mmHg)
- Baseline office SBP by tertiles and medians (mmHg)
- Baseline ambulatory heart rate by tertiles and medians (bpm)
- Baseline office heart rate by tertiles and medians (bpm)
- 24-Hour Pulse Pressure <60 vs. ≥60 mmHg (mmHg)
- Orthostatic hypertension at baseline yes vs. no
- Orthostatic tachycardia at baseline yes vs. no
- Baseline plasma renin activity <0.65 vs. ≥0.65 (ng/mL/h)
- Baseline aldosterone-renin ratio by tertiles
- Baseline aldosterone by tertiles (ng/dL)
- Number of ablations performed by tertiles (RDN arm only)
- Total number of ablations performed in branch vessels by tertiles (RDN arm only)
- Total number of ablations performed in main renal artery vessels by tertiles (RDN arm only)
- Total number of 45 second ablations performed by tertiles (RDN arm only)
- Medication adherent vs. non-adherent subjects at screening visit 2 and 6 months (from urine and serum tests)

- Accessory arteries treated yes vs. no
- 1 vs. 2 vs. 3 prescribed AH medication classes at baseline

9.12. COVID-19 Related Analyses

In accordance with FDA guidance document "FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic" [8], we will perform additional analyses to assess the effect of the COVID-19 pandemic on the study outcomes. Additional COVID-19 related analyses may be performed.

9.12.1. COVID-19 Timing Poolability Analysis

The primary endpoint for subjects can be classified as 1) taken before the COVID-19 impact; and 2) during the COVID-19 pandemic On 24MAR2020 a letter was sent to all study sites to suspend enrollment. Any 6-month follow-up visits performed before this date will be classified as "before COVID-19 impact", while those on or after this date will be classified as "during COVID-19 impact".The following poolability analysis will be performed betwen the COVID-19 impacted groups defined above.

- A linear regression will be performed, with change in 24-hour SBP from baseline to 6 months as the dependent variable, and baseline 24-Hour SBP, treatment, COVID-19 indicator, treatment * COVID-19 interaction as independent variables. If the interaction term is not significant at 0.15 level, then the primary efficacy endpoint results are considered consistent between COVID-19 subgroups.

9.12.2. COVID-19 Randomization Date Poolability Analysis

Any subjects randomized on 1AUG2019 or later will have their 6 month efficacy endpoint visit during the COVID-19 pandemic period. The following poolability analysis will be performed betwen subjects randomized before vs. on or after 1NOV2019.

- A linear regression will be performed, with change in 24-hour SBP from baseline to 6 months as the dependent variable, and baseline 24-Hour SBP, treatment, COVID-19 randomization date indicator, treatment * COVID-19 randomization date indicator interaction as independent variables. If the interaction term is not significant at 0.15 level, then the primary efficacy endpoint results are considered consistent between COVID-19 subgroups.

9.12.3. COVID-19 Positive vs. Negative Subgroup Analysis

We will identify which subjects test positive for COVID-19 and we will perform the following analysis comparing these subjects to those who do not test positive.

- A linear regression will be conducted, with 6-month change in 24-hour SBP as the dependent variable, and baseline 24-hour SBP, treatment, COVID-19 Yes/No, treatment * COVID-19 Yes/No interaction term as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary safety endpoint is considered consistent between subjects with and without COVID-19.

9.12.4. COVID-19 Method of Data Collection

An analysis will be performed to compare methods of data collection at 6-months; office visit, home visit, or remote follow-up visit (phone or virtual).

- A linear regression will be conducted, with 6-month change in 24-hour SBP as the dependent variable, and baseline 24-hour SBP, treatment, collection method (office/home/remote), treatment * collection method interaction term as independent variables. If the interaction term is not significant at 0.15 level, then the treatment effect in the primary safety endpoint is considered consistent between data collection methods.

9.12.5. Protocol Deviation and Adverse Event Reporting

In addition to the analyses described above, we will also provide COVID-19 related protocol deviation tables and AE/SAE tables summarizing the site-reported adverse events attributed to COVID-19 as recommended by FDA guidance document.

9.13. Changes to Planned Analysis

There are no changes to the planned analysis at this time.

10. Validation Requirements

Statistical programming for the analysis datasets, primary endpoints, secondary safety endpoints, and secondary effectiveness endpoints require Level 1 (independent) validation. Other objectives and sub-group analyses require Level 1 (independent) or Level 2 (Peer review) validation.

11. References

[1] http://mdic.org/computer-modeling/virtual-patients/ [Accessed: 8th May 2020]

[2] Haddad T, Himes A, Thompson L, Irony T, Nair R. Incorporation of stochastic engineering models as prior information in Bayesian medical device trials. *Journal of Biopharmaceutical Statistics*, 2017; 27(6): 1089-1103.

[3] Ibrahim JG, Chen MH. Power prior distributions for regression models, *Statistical Science,*

2000; 15(1): 46-60.

[4] Gelman A, Carlin JB, HS Ster, DB Rubin. *Bayesian Data Analysis*. Boca Raton: Chapman & Hall / CRC, 2004.

[5] Berry SM, Carlin BP, Lee JJ, Muller P. *Bayesian Adaptive Methods for Clinical Trials*. Boca Raton: CRC Press, 2010.

[6] Kandzari DE, Böhm M, Mahfoud F, Townsend RR, Weber MA, Pocock S, Tsioufis K, Tousoulis D, Choi JW, East C, Brar S, Cohen S, Fahy M, Pilcher G, Kauomi K. Effect of renal denervation on blood pressure in the presence of antihypertenisve drugs: 6-month efficacy and safety results from the SPYRAL HTN-ON MED proof of concept randomized trial. *The Lanc*et, 2015; 391(10137): 2346-2355.

[7] Sanders, Margreet F., et al. Renal artery and parenchymal changes after renal denervation: assessment by magnetic resonance angiography. *European Radiology* 2017; 27(9): 3934-3941.

[8] FDA Guidance on Conduct of Clinical Trials of Medical Products during COVID-19 Pandemic. Guidance for Industry, Investigators, and Institutional Review Boards. March 2020.

[9] Chen, Wei Chen, Chenguang Wang, Heng Li, Nelson Lu, Ram Tiwari, Yunling Xu, and Lilly Q. Yue. 2020. "Propensity score-integrated composite likelihood approach for augmenting the control arm of a randomized controlled trial by incorporating real-world data." Journal of Biopharmaceutical Statistics 30 (3): 508–20. https://doi.org/10.1080/10543406.2020.1730877

[10] Wang, Chenguang, Heng Li, Wei Chen Chen, Nelson Lu, Ram Tiwari, Yunling Xu, and Lilly Q Yue. 2019. "Propensity score-integrated power prior approach for incorporating real-world evidence in single-arm clinical studies." Journal of Biopharmaceutical Statistics 29 (5): 731–48. https://doi.org/10.1080/10543406.2019. 1657133.

[11] Wang, Chenguang, Nelson Lu, Wei Chen Chen, Heng Li, Ram Tiwari, Yunling Xu, and Lilly Q. Yue. 2020. "Propensity score-integrated composite likelihood approach for incorporating real-world evidence in single arm clinical studies." Journal of Biopharmaceutical Statistics 30 (3): 495–507. https://doi.org/10.1080/ 10543406.2019.1684309.

[12] Hobbs, B. P., Carlin, B. P., Mandrekar, S. J. 2011. "Hierarchical commensurate and power prior models for adaptive incorporation of prior information in clinical trials." *Biometrics* 67, 1047–1056.

[13] Pastore M (2018). "Overlapping: a R package for Estimating Overlapping in Empirical Distributions." The Journal of Open Source Software, 3(32), 1023. https://doi.org/10.21105/joss.01023)

[14] D. Kandzari, G. Hickey, S. Pocock, "Prioritized Endpoints for Device-Based Hypertension Trials: the Win Ratio Methodology", EuroIntervention, DOI: 10.4244/EIJ-D-20-01090.

[15] AC Flint, DL Bhatt, et al. "Effect of Systolic and Diastolic Blood Pressure on Cardiovascular Outcomes", N Engl J Med 2019;381:243-51. DOI: 10.1056/NEJMoa1803180

[16] N Fatani, DL Dixon, et al. "Systolic Blood Pressure Time in Target Range and Cardiovascular Outcomes in Patients With Hypertension", J Am Coll Cardiol 2021;77:1290–9, https://doi.org/10.1016/j.jacc.2021.01.014

12. Statistical Appendices

12.1. Appendix I: Imputation of Missing Dates

Imputation of Missing AE Onset Date

Valid Portion	Missing Portion	Imputed Value for Missing Portion
Month, Year	Day	Set Day = first day of that month and year, then set the day = later of (New onset date, informed consent date).
Year	Day, Month	Set date = later of (January 1st of that year, informed consent date).
None	Day, Month, Year	Informed consent date.

Imputation of Missing Medication Start Date

Valid Portion	Missing Portion	Imputed Value for missing Portion
Month, Year	Day	Set Day = first day of that month and year
Year	Day, Month	Set date = January 1st of that year
None	Day, Month, Year	SV2 date

Imputation of Missing Medication Stop Date

Valid Portion	Missing Portion	Imputed Value for missing Portion
Month, Year	Day	Set Day = first day of that month and year, then set the day = later of (New date, SV2 date, start date).
Year	Day, Month	Set date = later of (January 1st of that year, SV2 date, start date).