


Transcatheter Aortic Valve Replacement With the Medtronic Transcatheter Aortic Valve Replacement System In Patients at Low Risk for Surgical Aortic Valve Replacement

Clinical Trials.gov Identifier: NCT02701283

Statistical Analysis Plan

Document Date: 17 JUL 2018 (Version 4.0)

 <b>Statistical Analysis Plan</b>	
<b>Clinical Investigation Plan Title</b>	Transcatheter Aortic Valve Replacement With the Medtronic Transcatheter Aortic Valve Replacement System In Patients at Low Risk for Surgical Aortic Valve Replacement
<b>Clinical Investigation Plan Identifier</b>	10234430DOC
<b>Sponsor/Local Sponsor</b>	Medtronic, Inc. Coronary and Structural Heart Clinical 8200 Coral Sea St NE, MVS 66 Mounds View, MN 55112 United States
<b>Document Version</b>	4.0
<b>Confidentiality Statement</b>	
<p>The information contained in this document is confidential and the proprietary property of Medtronic. Any distribution, copying, or disclosure without the prior written authorization of Medtronic is strictly prohibited. Persons to whom the information is disclosed must know that it is confidential and that it may not be further disclosed by them.</p>	

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

## Table of Contents

<b>1.</b>	<b>Version History .....</b>	<b>4</b>
<b>2.</b>	<b>Introduction .....</b>	<b>4</b>
<b>3.</b>	<b>Study Objectives.....</b>	<b>5</b>
3.1.	Primary Safety and Effectiveness Endpoint .....	5
3.2.	Secondary Safety Endpoints .....	5
3.3.	Secondary Effectiveness Endpoints .....	5
<b>4.</b>	<b>Investigation Plan .....</b>	<b>6</b>
<b>5.</b>	<b>Determination of Sample Size .....</b>	<b>7</b>
5.1.	Historical Data .....	7
5.2.	Sample Size .....	7
<b>6.</b>	<b>Statistical Methods.....</b>	<b>8</b>
6.1.	Randomization .....	8
6.2.	Analysis Populations .....	8
6.2.1.	Screening Population .....	8
6.2.2.	Randomized Population .....	8
6.3.	Description of Baseline Variables .....	9
6.4.	Kaplan-Meier Analyses.....	9
6.5.	Primary Objective .....	9
6.5.1.	Hypothesis of non-inferiority .....	9
6.5.2.	Analysis Plan .....	10
6.5.3.	Note on Timing of Stroke Determination .....	10
6.6.	Description of Performed Analysis, per Population.....	11
6.7.	Secondary Safety Endpoints.....	11
6.8.	Secondary Effectiveness Endpoints .....	11
6.9.	Multiplicity Considerations.....	13
6.9.1.	Ordered List of Secondary Objectives to be Tested to Support Labeling Claims .....	14
6.10.	Heterogeneity/Poolability .....	15
6.10.1.	Geography Poolability Analysis.....	15
6.10.1.1.	Primary Endpoint by Geography .....	15
6.10.1.2.	Univariate Covariate Analysis .....	15
6.10.1.3.	Multivariable Analysis .....	16
6.10.2.	Site Poolability Analysis .....	16

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

6.10.2.1.	Pooling of Small Sites.....	16
6.10.2.2.	Primary Endpoint by Site.....	16
6.10.2.3.	Multivariable Analysis.....	17
<b>7.</b>	<b>Additional Analysis .....</b>	<b>17</b>
<b>8.</b>	<b>Validation Requirements .....</b>	<b>17</b>
<b>9.</b>	<b>References .....</b>	<b>17</b>

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

## 1. Version History

Version	Summary of Changes	Author(s)/Title
1.0	Initial release	Hongyan Qiao Principal Statistician
2.0	Added details for the Bayesian study design; Incorporated the CIP updates.	Hongyan Qiao, PhD Sr. Principal Statistician Andrew S Mugglin, PhD Paradigm Biostatistics, LLC
3.0	Change the first interim analysis timing to 850 subjects have had the chance to finish 1 year follow up; Delete “the non-parametric Wilcoxon rank-sum test”.	Hongyan Qiao, PhD Sr. Principal Statistician Andrew S Mugglin, PhD Paradigm Biostatistics, LLC
4.0	Add clarification for the primary endpoint superiority testing order; Add Section 6.9 Multiplicity Considerations; Rename “additional outcome measures” to “secondary effectiveness endpoints”.	Hongyan Qiao, PhD Sr. Principal Statistician

## 2. Introduction

This Statistical Analysis Plan has been designed to document, before data are analyzed, the rationale for the study design, and the planned analyses that will be included in study reports. This Statistical Analysis Plan (SAP) is developed based on the Transcatheter Aortic Valve Replacement (TAVR) Low Risk Clinical Investigational Plan (CIP).

The purpose of this trial is to evaluate the safety and effectiveness of transcatheter aortic valve replacement (TAVR) in the treatment of symptomatic severe aortic stenosis (AS) in subjects who are determined by the Heart Team to be at low surgical risk. The Low Risk trial is a multicenter, prospective, 1:1 randomized study designed to demonstrate non-inferiority (within an absolute margin of 6%) of TAVR to SAVR, as measured by a composite of all-cause death or disabling stroke rate at 24 months. The planned sample size is 1200 subjects who undergo an attempted study procedure, and the statistical methods are Bayesian. Study success will be evaluated according to an analysis plan that includes two possible interim analyses, as well as a final analysis. The first interim analysis is timed when 850 subjects have had the chance to be followed for 12 months (12 months after the 850<sup>th</sup> procedure date); a second interim analysis is timed when 1200 subjects have had the chance to be followed for 12 months (12 months after the 1200<sup>th</sup> procedure date); and the final analysis is timed when all subjects have had the chance to be followed for 24 months (24 months after the last LTI subject’s procedure date).

A study report will be prepared for submission to US Food and Drug Administration (FDA) at the time study success criteria have been met, for the purpose of seeking market approval. After all subjects have completed all protocol-specified follow-up, a final clinical report including updated long-term safety data will be prepared and submitted.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

Any deviations from the Statistical Analysis Plan will be described and justified in the Final Report, as appropriate.

### **3. Study Objectives**

The primary objective of the trial is to demonstrate that the safety and effectiveness of the Medtronic TAVR system, as measured by the rate of all-cause mortality or disabling stroke at 2 years, is non-inferior to SAVR in the treatment of severe aortic stenosis in subjects who have a low predicted risk of mortality at 30 days for SAVR. The following endpoints will be used to evaluate the primary trial objectives:

#### **3.1. Primary Safety and Effectiveness Endpoint**

The rate of all-cause mortality or disabling stroke at 2 years

#### **3.2. Secondary Safety Endpoints**

- The rate of the composite of death, disabling stroke, life-threatening bleed, major vascular complication, or AKI (II or III) at 30 days
- The rate of new permanent pacemaker implantation at 30 days
- The rate of prosthetic valve endocarditis at one year
- The rate of prosthetic valve thrombosis at one year
- The rate of all stroke (disabling and non-disabling) at one year
- The rate of life-threatening bleeding at one year
- The rate of valve-related dysfunction requiring repeat procedure at one year

#### **3.3. Secondary Effectiveness Endpoints**

- The rate of valve-related dysfunction, defined as moderate or severe prosthetic valve stenosis, or moderate or severe prosthetic regurgitation at one year (per VARC II)
- Quality of Life as assessed by Kansas City Cardiomyopathy Questionnaire (KCCQ) at 30 days and one year
- The rate of repeat hospitalization for aortic valve disease at one year
- Device Success (VARC II), defined as
  - Absence of procedural mortality, AND
  - Correct positioning of a single prosthetic heart valve into the proper anatomical location, AND
  - Intended performance of the prosthetic heart valve, defined as the absence of patient-prosthesis-mismatch and mean aortic valve gradient less than 20 mmHg (or peak velocity < 3 m/sec), AND absence of moderate or severe prosthetic valve regurgitation.
- Hemodynamic performance metrics by Doppler echocardiography
  - Mean aortic gradient at one year
  - Effective orifice area at one year
  - Degree of total, peri, and transvalvular prosthetic regurgitation at one year
- New York Heart Association (NYHA) functional classification at one year
- Health-related quality of life at one year as assessed by EQ-5D survey instrument

This document is electronically controlled. Printed copies are considered uncontrolled.

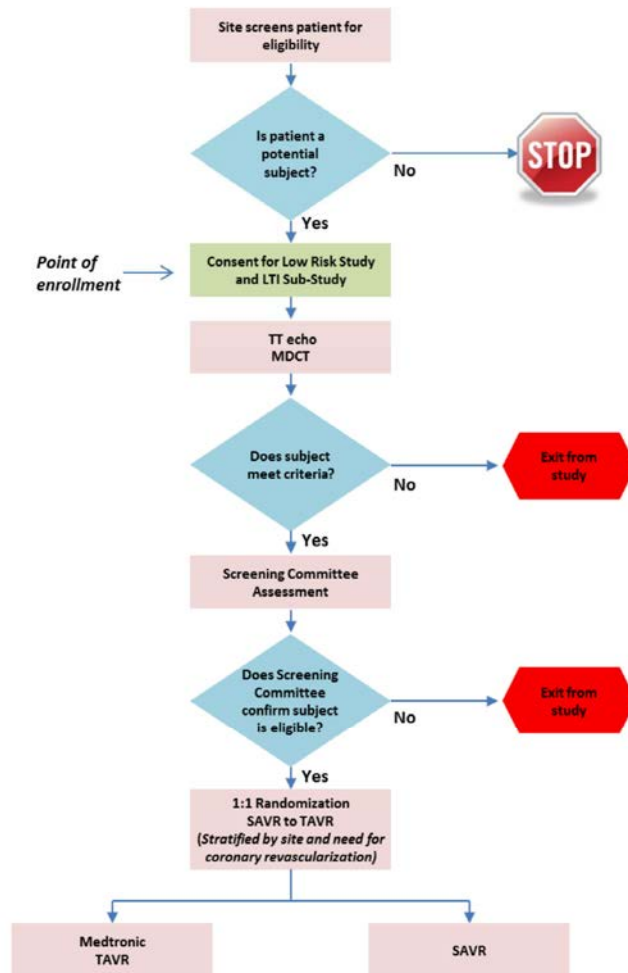
**Confidential**

## 4. Investigation Plan

No site will implant more than 100 subjects without prior authorization from Medtronic. Subjects who exit from the trial after implantation will not be replaced.

Subjects will be consented for follow-up through ten years. The enrollment period is estimated to be between 18 to 24 months; therefore the estimated total duration of the trial (first subject enrolled to last subject completing his/her last follow-up exam) is estimated to be twelve years.

The process of patient screening, subject enrollment, and randomization is as follows:



### Notes

1. TTE, MDCT, coronary arteriography, or labs performed for diagnostic purposes prior to consent may be used for the baseline/screening exams, provided they were performed within window and contain the necessary data.
2. Only the sites participating in the LTI Sub-study will consent for both the main study and LTI Sub-study.
3. Subjects who give consent for LTI Sub-study will follow sub-study protocol in addition to main protocol.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

## 5. Determination of Sample Size

### 5.1. Historical Data

Although the experience with surgical aortic valves in low surgical risk aortic valve replacement populations is extensive, it was not considered possible to leverage much of the data from these studies directly as surgical aortic valve replacement series typically enroll patient populations which include a proportion subjects with a bicuspid or unicuspid valve (excluded from this trial) which may be as high as approximately 50% and also include subjects with purely or primarily regurgitant lesions (also excluded from this trial) whose outcomes may differ from patients with aortic stenosis. As a result, series for which individual patient data were available or which were known to attempt exclusion of patients with bicuspid or unicuspid valves comprise the primary basis for the event rate estimate and additional series were considered only confirmatory in nature. Table 1 presents the rates of all-cause mortality at 24 months from studies which were considered in developing the event rate estimate. The simple weighted average from the studies was 11.4% which was adjusted up to 12% to account for the possibility that surgical candidates at extremely low risk (i.e. the healthiest and youngest potential subjects) may forego randomization into a TAVR trial until longer term data and data from lower surgical risk patients are available for TAVR.

**Table 1. Surgical series supporting all-cause mortality event rate estimate**

Surgical Series	Number in Cohort	24-Month All-cause Mortality K-M Rate
NOTION SAVR cohort <sup>1</sup>	134	9.8%
3f Pivotal cohort <sup>1</sup>	405	12.7%
Mosaic sub-analysis <sup>2</sup>	646	9.9%
Freestyle sub-analysis <sup>3</sup>	323	13.5%
Simple Weighted Average	1508	11.4%
<sup>1</sup> Entire cohort leveraged (data on file) <sup>2</sup> Sub-group analysis excluding subjects <65 years of age, with congenital bicuspid valves or with a purely regurgitant lesion <sup>3</sup> Sub-group analysis excluding subjects <65 years of age, with congenital bicuspid valves, with a purely regurgitant lesion or with the Freestyle valve implanted as a full-root replacement		

The 24-month event rate estimate for non-fatal disabling stroke of 3% was generated primarily from the CoreValve US Pivotal High Risk Trial SAVR (3.9%) and TAVR (2.0%) cohorts which used event definitions consistent with this trial.<sup>1</sup> Additional sources of data considered include the NOTION SAVR cohort (4.6%) and the Mosaic and Freestyle cohort sub-analyses (which had rates of 2.0% and 2.3% respectively) all of which collected clinical stroke/cerebrovascular accidents.

The incidence of the primary endpoint of all-cause mortality or disabling stroke at 24 months in each treatment group is expected to be 15%, which is based on an assumed rate of all-cause mortality of 12% at 24 months and a non-fatal disabling stroke rate of 3% at 24 months.

### 5.2. Sample Size

Although the pre-specified analysis methods are Bayesian, the sample size is guided by a standard frequentist non-inferiority power analysis. Under the assumptions of the 15% incidence in each treatment group, non-inferiority margin  $\delta=0.06$ , 1:1 randomization,  $\alpha=0.05$ , and power = 85%, the method of Farrington and Manning<sup>2</sup> as implemented in PASS 2013<sup>3</sup> indicates that the required sample size for a single-look analysis is 1032. To allow for up to 6% dropout, 1100 subjects must be accrued. Furthermore, to compensate for power lost in a three-look group

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**



sequential analysis plan using equal information increments and Pocock-type alpha spending, the sample size would have to be increased by about 15%, whereas an O'Brien-Fleming alpha spending approach would increase the total sample size by approximately 1%<sup>4</sup>. The impact of the sampling plan used in this Bayesian design is likely to fall between these two approaches, so an additional 100 subjects (9%) will be accrued, bringing the total estimated sample size to 1200 as treated (AT) subjects, which should provide ample power for establishing non-inferiority in the primary hypothesis test.

## 6. Statistical Methods

### 6.1. Randomization

Randomization will follow a 1:1 allocation ratio and be stratified by site and need for revascularization, using a blocked randomization scheme with blocks of randomly varying sizes.

### 6.2. Analysis Populations

#### 6.2.1. Screening Population

All patients with symptomatic severe AS who provide informed consent will be considered screened and all available data will be entered into the Electronic Data Capture (EDC) system.

#### 6.2.2. Randomized Population

If the subject signs informed consent, meets all inclusion and none of the exclusion criteria, and the Heart Team determines the subject is suitable for randomization in the trial, then the subject is reviewed by the Screening Committee. If the subject is approved by the Screening Committee and the subject is randomized to either TAVR or SAVR, the subject is added to the randomized population. Within the randomized population the following analysis sets are distinguished:

- **The intention to treat (ITT) set:** Subjects are reported according to the randomized assignment, SAVR or TAVR, regardless of what, if any, therapy was actually received. Time zero begins at the date of randomization.
- **The as treated (AT) set:** The AT set consists of all ITT subjects with an attempted implant procedure, defined as when the subject is brought into the procedure room and any of the following have occurred: anesthesia administered, vascular line placed, TEE placed or any monitoring line placed. Subjects will be analyzed according to their first attempted procedure (TAVR or SAVR). Time zero begins at the date of the first TAVR or SAVR attempted procedure.
- **The implanted set:** The Implanted set consists of all the AT subjects who are actually implanted with either the Transcatheter Aortic Valve (TAV) or Surgical Aortic Valve (SAV). Time zero begins at the date of the first TAVR or SAVR attempted procedure.
- **The per protocol (PP) set:** The PP set is defined based on the International Council for Harmonisation (ICH) E9 Statistical Principles, which will consist of the following:
  1. All implanted subjects who were implanted according to their randomization; and
  2. Subjects without early exit (eg., lost of follow up) before 24 months (730 days), except those experiencing the primary endpoint (death or disabling stroke) prior to the early exit; and

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

3. Subjects without crossover to a different type of procedure from their first attempted procedure type before their 24-month visits; and
4. Subjects must satisfy all inclusion/exclusion criteria.

Time zero begins at the date of the first TAVR or SAVR attempted procedure.

The primary analysis for the primary objective, secondary safety objectives, secondary effectiveness objectives (except for valve dysfunction, hemodynamic performance metrics, and device success) will use the AT set. Valve dysfunction, hemodynamic performance metrics, and device success will use the implanted set.

### 6.3. Description of Baseline Variables

Baseline demographic and clinical variables will be summarized for each of the treatment groups for the intention-to-treat (ITT), as treated (AT), and implanted sets. All continuous variables will be summarized as means, medians, standard deviations, interquartile ranges, minima and maxima and will be compared between treatment groups via 95% Bayesian credible intervals (BCIs) for the difference in means. Categorical variables will be summarized as frequencies and percentages and will be compared between treatment groups via 95% Bayesian credible intervals for the difference in proportions.

### 6.4. Kaplan-Meier Analyses

For safety related endpoints, the Kaplan-Meier event rates at 30 days, 6 months, 12 months, 18 months and annually through 10 years will be provided. For these analyses, these times correspond to 30 days, 183 days, 365 days, 545 days, 730 days, and annually through 10 years (365×3, 365×4, 365×5, etc.). At each time point with data, the product-limit estimate of the event rate, the number of subjects at risk, the number of subjects with events, the Peto standard error of the estimate, and the loglog transformed 95% confidence interval using the Peto standard error will be presented.

For subjects without an event, the date of censoring will be the latest date of all follow-up visits, assessments, and events (including death in those endpoints where death is not the endpoint).

### 6.5. Primary Objective

The primary endpoint of all-cause mortality or disabling stroke at 24 months will be evaluated using the absolute difference of the TAVR rate and the SAVR rate for all-cause mortality or disabling stroke during a fixed follow-up of 24 months. The hypothesis test is designed to show non-inferiority of TAVR to SAVR for the primary endpoint.

#### 6.5.1. Hypothesis of non-inferiority

The primary objective is to establish that TAVR is non-inferior to SAVR for the primary endpoint. The hypothesis of interest is:

$$H: \pi_T < \pi_C + \delta$$

where  $\pi_T$  and  $\pi_C$  respectively denote incidence of all-cause mortality or disabling stroke at 24 months for the treatment (TAVR) and control (SAVR) groups, and  $\delta = 0.06$ . This study is designed using Bayesian statistical techniques. TAVR will be considered to be non-inferior to SAVR if it can be established that the posterior probability  $\Pr(H_{\delta=0.06} | \text{data}) > \Psi$ , where  $\Psi$  is a pre-specified threshold value. In addition, the primary endpoint (superiority) will be tested according to the testing order specified in Section 6.9. The values chosen for  $\Psi$  and  $\Psi_{\text{SUP}}$  are chosen to ensure type I error rates are not larger than 0.05 for non-inferiority and 0.025 for superiority.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

### 6.5.2. Analysis Plan

The two interim analyses and final analysis are planned when:

1. 850 subjects have had the chance to be followed for 12 months (12 months after the 850<sup>th</sup> procedure date);
2. 1200 subjects have had the chance to be followed for 12 months (12 months after the 1200<sup>th</sup> procedure date);
3. All subjects have had the chance to be followed for 24 months (24 months after the last LTI subject's procedure date).

At the first interim analysis, the posterior probability of non-inferiority  $P(H_{\delta=0.06} | \text{data})$  will be calculated, and if this probability exceeds a pre-specified threshold  $\Psi$ , non-inferiority will be concluded. This will be considered an “early win” for non-inferiority, and a regulatory submission will follow. Otherwise, follow-up will continue until the second interim analysis, where again the posterior probability of non-inferiority  $P(H_{\delta=0.06} | \text{data})$  will be calculated, and if this probability exceeds a pre-specified threshold  $\Psi$ , non-inferiority will be concluded and a regulatory submission will follow. If non-inferiority is not concluded at either of the interim analyses, follow-up will continue until all subjects have had the chance to be followed for 24 months (24 months after the last LTI subject's procedure date), and non-inferiority will be tested a third time. The standard for concluding non-inferiority in each case is the same:  $P(H_{\delta=0.06} | \text{data}) > \Psi$ .

If non-inferiority is established at either interim analysis, a test of superiority may be performed (See Section 6.9 for additional requirements and testing sequence). If  $P(H_{\delta=0} | \text{data}) > \Psi_{\text{SUP}}$ , superiority will be established at this time. However, if  $P(H_{\delta=0} | \text{data}) \leq \Psi_{\text{SUP}}$ , subjects will continue to be followed until the full cohort has had the chance to be followed for 24 months (24 months after the last LTI subject's procedure date), at which time a delayed determination of superiority may be made if  $P(H_{\delta=0} | \text{data}) > \Psi_{\text{SUP}}$ .

If non-inferiority is not established until the final analysis, then superiority will only be tested at that time per the testing sequence outlined in Section 6.9.

The maximum number of non-inferiority assessments is three (first interim analysis, second interim analysis, final analysis). The maximum number of superiority assessments is two (simultaneous to passing the non-inferiority assessment for the primary endpoint and the hierarchical testing for the secondary endpoints listed in Section 6.9, or at the final analysis in the event that superiority is not established at one of the interim analyses).

The statistical approach for these analyses is Bayesian. The prior distributions for  $\pi_T$  and  $\pi_C$  in these calculations are Beta(1,1). In each analysis (interim or final), any subject with missing 24-month outcome will have that outcome predicted via a Bayesian piecewise exponential survival model that incorporates follow-up to date. Within a treatment group, let a subject's time to event follow a piecewise exponential model whose hazard function  $\lambda(\tau)$  is piecewise constant over the 3 partitioning intervals defined by the cutoff points 0, 30 days, 183 days, and 730 days. The parameters for each of the 3 intervals are assigned diffuse Gamma prior distributions, with mean 1 and variance 100, which at the time of analysis are updated based on observed data, and unobserved 24-month outcomes are predicted with probabilities determined by this model. Combining predicted and observed 24-month outcomes and integrating out the predictive distributions results in a posterior probability of non-inferiority (or superiority) that accounts for missing data as well as the uncertainty in the prediction, and this is the quantity that is compared to  $\Psi$  (or  $\Psi_{\text{SUP}}$ ) to assess non-inferiority (or superiority). This approach is similar to that taken in the SURTAVI trial<sup>5</sup>, though in this case the statistical model underlying the predictions is based on a time-to-event model rather than a set of beta-binomial models. It is also similar to the approach taken in Wilber et al<sup>6</sup>. The thresholds  $\Psi$  and  $\Psi_{\text{SUP}}$  are selected empirically to achieve a type I error rate (under extensive simulation) of at most 0.05 for non-inferiority testing and at most 0.025 for superiority testing.

### 6.5.3. Note on Timing of Stroke Determination

It may take up to 90 days after a stroke is reported to determine whether it is a disabling stroke. In such cases, the date of the stroke is the date of occurrence and not the date of determination. Furthermore, when the database is locked for interim and final analyses, it is possible that some strokes will still be waiting for the 90-day MRS. In such cases,

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

the CEC will make their adjudications based on all available data, and the analysis will be based on the CEC adjudicated results.

## 6.6. Description of Performed Analysis, per Population

The primary analysis for the primary endpoint will be performed on the AT set. In addition, it will also be performed on the ITT, Implanted, and PP sets. The inferential statistics for the following secondary endpoints will be performed on the AT set:

- The rate of the composite of death, disabling stroke, life-threatening bleed, major vascular complications, or AKI (II or III) at 30 days
- The rate of new permanent pacemaker implantation at 30 days
- The rate of prosthetic valve endocarditis at one year
- The rate of prosthetic valve thrombosis at one year
- The rate of all stroke (disabling and non-disabling) at one year
- The rate of life-threatening bleeding at one year
- The rate of valve-related dysfunction requiring repeat procedure at one year
- Quality of Life as assessed by Kansas City Cardiomyopathy Questionnaire (KCCQ) at one year
- The rate of repeat hospitalization for aortic valve disease at one year
- New York Heart Association (NYHA) functional classification at 30 days, 6 months, one year, 18 months, 2 years, 3 years, 4 years, 5 years, 7 years and 10 years
- Health-related quality of life at one year as assessed by EQ-5D survey instrument

The implanted set will be used for analyzing the primary endpoint, secondary endpoint of prosthetic valve dysfunction, device success, and echocardiographic assessment of valve performance.

## 6.7. Secondary Safety Endpoints

The following secondary endpoints will be compared between TAVR and SAVR subject cohorts using the appropriate Bayesian version of analysis for comparing proportions. In addition, Kaplan-Meier estimates will be provided.

- The rate of the composite of death, disabling stroke, life-threatening bleed, major vascular complications, or AKI (II or III) at 30 days
- The rate of new permanent pacemaker implantation at 30 days
- The rate of prosthetic valve endocarditis at one year
- The rate of prosthetic valve thrombosis at one year
- The rate of all stroke (disabling and non-disabling) at one year
- The rate of life-threatening bleeding at one year
- The rate of valve-related dysfunction requiring repeat procedure at one year

## 6.8. Secondary Effectiveness Endpoints

- The rate of valve-related dysfunction, defined as moderate or severe prosthetic valve stenosis, or moderate or severe prosthetic regurgitation at one year (per VARC II)

The incidence estimate will be provided for the two treatment groups at the specified time point. The statistical method will be the Bayesian version of a comparison of proportions. The incidence estimates will also be reported at the following time points: 30 days, 6 months, one year, 2 years, 3 years, 4 years, 5 years, 7 years and 10 years.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

<b>PROSTHETIC VALVE DYSFUNCTION</b>	
Stenosis: moderate/severe	Any of the following 1) Peak aortic velocity >4 m/s OR mean aortic gradient >40 mmHg, AND EOA <0.8 cm <sup>2</sup> . 2) Peak aortic velocity >4 m/s OR mean aortic gradient >40 mmHg, AND EOA ≥0.8 cm <sup>2</sup> , and DVI <0.25, 3) Peak aortic velocity ≤4 m/s and mean aortic gradient ≤ 40 mmHg, AND EOA <0.8 cm <sup>2</sup> , and DVI <0.25
Paravalvular regurgitation	Moderate/ Severe paravalvular regurgitation
Transvalvular regurgitation	Moderate/ Severe transvalvular regurgitation
Total regurgitation	Moderate/ Severe total regurgitation

## Notes:

1. DVI = Doppler Velocity Index (LVOT VTI/valve VTI)
  2. For subjects with BSA <1.6 m<sup>2</sup>, the EOA criteria for significant (moderate or severe) stenosis is <0.6 cm<sup>2</sup>
  3. For subjects with LVOT diameter >2.5 cm, the DVI criteria for significant (moderate or severe) stenosis is <0.2 cm<sup>2</sup>
  4. Reporting of prosthetic valve dysfunction will be based on core lab results.
  5. Prosthetic valve dysfunction events are not reported as adverse events, unless the dysfunction is accompanied with clinical sequelae at the time of event detection, and the clinical sequelae are chronologically and physiologically associated with the dysfunction. However, prosthetic dysfunctions that are associated with adverse events, and that meet the definition of a serious adverse event, should be reported as such.
- Quality of Life as assessed by Kansas City Cardiomyopathy Questionnaire (KCCQ) at 30 days and one year  
The endpoint will be evaluated using a Bayesian analog of a two-sample t-test. Descriptive statistics for KCCQ change from baseline will also be reported at 30 days, 6 months, one year, and annually through 5 years.
  - The rate of repeat hospitalization for aortic valve disease at one year.  
The incidence estimate will be provided for the two treatment groups at the specified time point. The statistical method will be the Bayesian version of analysis for comparing proportions. In addition, the Kaplan-Meier estimates will be provided at the following time points: 30 days, 6 months, 12 months, 18 months, and annually thereafter through 10 years.
  - Device Success (VARC II), defined as
    - Absence of procedural mortality, AND
    - Correct positioning of a single prosthetic heart valve into the proper anatomical location, AND
    - Intended performance of the prosthetic heart valve, defined as the absence of patient-prosthesis-mismatch and mean aortic valve gradient less than 20 mmHg (or peak velocity < 3 m/sec), AND absence of moderate or severe prosthetic valve regurgitation.

The incidence estimate will be provided for the TAVR group.

The criteria for ECHO will be based on the ECHO Core Lab Data:

- Absence of patient-prosthesis-mismatch
  1. For subjects with BMI < 30 kg/m<sup>2</sup>, index effective orifice area (EOAi) > 0.85 cm<sup>2</sup> /m<sup>2</sup>
  2. For subjects with BMI ≥ 30 kg/m<sup>2</sup>, index effective orifice area (EOAi) > 0.70 cm<sup>2</sup> /m<sup>2</sup>
  3. BMI = weight(kg)/(height (m))<sup>2</sup>

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

- Mean aortic gradient < 20 mmHg or peak velocity < 3 m/sec;
- Absence of moderate or severe prosthetic valve regurgitation. Total Aortic Prosthetic Regurgitation on the ECHO CORE LAB form not equal to moderate or severe.

All of the above components must be satisfied to count as a device success. If any of the above components fails, the endpoint will be counted as a failure.

For the overall device success rate, the numerator will be the number of subjects whose procedures result in device success as described above, and the denominator will be the number of subjects whose device success results are not missing (either success or failure). Note that this analysis excludes those subjects with a missing response to any of the above five components (eg, the field “Post-implant Severity of Total Aortic Regurgitation”=“Unable to Assess” or “Not Recorded”, missing mean aortic gradient, or missing peak velocity, etc.) and without a “NO” response to any of the components.

- Hemodynamic performance metrics by Doppler echocardiography
  - Mean aortic gradient at one year
  - Effective orifice area at one year
  - Degree of total, peri, and transvalvular prosthetic regurgitation at one year

For mean gradient and effective orifice area, the descriptive statistics will be reported at each of the assessed time point (30 days, 6 months, one year, 2 years, 3 years, 4 years, 5 years, 7 years and 10 years). The statistical method will be the Bayesian version of a two-sample t-test.

Prosthetic regurgitation severity will be reported as proportions at each of the assessed time point (30 days, 6 months, one year, 2 years, 3 years, 4 years, 5 years, 7 years and 10 years). The statistical method will be the Bayesian version of a comparison of proportions.

- New York Heart Association (NYHA) functional classification at one year

NYHA function classification will be reported as proportions at each of the assessed time point (30 days, 6 months, one year, 2 years, 3 years, 4 years, 5 years, 7 years and 10 years). The statistical method will be the Bayesian version of a comparison of proportions.

- Health-related quality of life as assessed by EQ-5D survey instrument at one year.

The endpoint will be evaluated using a Bayesian analog of a two-sample t-test.

## 6.9. Multiplicity Considerations

It is recognized that with a multiplicity of tests comes an inflation in the chance of a false finding of superiority or non-inferiority. Therefore, for the purpose of seeking approved labeling claims on designated secondary objectives, the following standard will be used: if the primary objective demonstrates non-inferiority, claims will be sought for selected secondary non-inferiority and superiority objectives and for superiority on the primary objective metric. These will be tested via a hierarchical (sequential) testing order that preserves the overall study-wise type I error rate at the level of 0.05. The testing order is specified below. The following secondary objectives are tested in order, and testing continues if and only if all previous objectives have met their designated success criterion.

1. Transvalvular mean gradient at 1 year (non-inferiority)
2. Effective orifice area at 1 year (non-inferiority)
3. Change in NYHA classification from baseline to 1 year (non-inferiority)
4. Change in KCCQ score from baseline to 1 year (non-inferiority)
5. Transvalvular mean gradient at 1 year (superiority)
6. Effective orifice area at 1 year (superiority)
7. Change in KCCQ score from baseline to 30 days (superiority)

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

All of the above non-inferiority will be tested with a type I error standard of 0.05 and superiority tests will be tested with a type I error standard of 0.025. If all of the above tests meet their success criterion, the primary endpoint (superiority) will be tested using a type I error rate of 0.025.

For the purposes of seeking claims, these objectives will only be evaluated once, at the same time as non-inferiority of the primary objective is established. The only exception to this is the primary endpoint superiority test, which carries the possibility of a delayed determination of superiority and may thus meet its success criterion at a different time (see Section 6.5.2).

The remaining secondary endpoints (see Section 3.2 and Section 3.3) may be of interest for scientific reasons but will not be the basis for supporting labeling claims; they are thus outside of the hierarchical testing procedure. Similarly, for those objectives that test non-inferiority, if non-inferiority is established, a test of superiority may also be conducted, but unless specifically itemized in the list, such superiority testing is not part of the hierarchical testing procedure; these superiority tests may be of interest for scientific reasons but will not be the basis for supporting labeling claims.

### 6.9.1. Ordered List of Secondary Objectives to be Tested to Support Labeling Claims

1. Transvalvular mean gradient at 1 year (non-inferiority). The hypothesis of interest is

$$H: \mu_{TAVR} < \mu_{SAVR} + 5$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the average mean gradient at 1 year, measured in mmHg. This objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.95.

**Rationale for Delta:** A difference less than 5 mmHg for mean gradient is not considered clinically relevant for the TAVR Low Risk Executive Committee.

2. Effective orifice area at 1 year (non-inferiority). The hypothesis of interest is

$$H: \mu_{TAVR} > \mu_{SAVR} - 0.1$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the mean effective orifice area at 1 year, measured in  $\text{cm}^2$ . This objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.95.

**Rationale for Delta:** A difference less than 0.1  $\text{cm}^2$  for effective orifice area is not considered clinically relevant for the TAVR Low Risk Executive Committee.

3. Change in NYHA classification from baseline to 1 year (non-inferiority). The hypothesis of interest is

$$H: \mu_{TAVR} > \mu_{SAVR} - 0.375$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the mean number of classification improvements in NYHA from baseline to 1 year. For subjects with NYHA categories at both baseline and 1-year visit, the NYHA classification improvements will be calculated as  $\text{NYHA}_{\text{baseline}} - \text{NYHA}_{12\text{month}}$ . The objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.95.

**Rationale for Delta:** A difference less than 0.375 for NYHA classification is not considered clinically relevant for the TAVR Low Risk Executive Committee.

4. Change in Kansas City Cardiomyopathy Questionnaire (KCCQ) score from baseline to 1 year (non-inferiority). The hypothesis of interest is

$$H: \mu_{TAVR} > \mu_{SAVR} - 5$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the mean changes in the KCCQ score from baseline to 1 year. For subjects with KCCQ score at both baseline and 1 year, the change in KCCQ will be calculated as  $\text{KCCQ}_{1\text{year}} - \text{KCCQ}_{\text{baseline}}$ . The objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.95.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

**Rationale for Delta:** A 5-point improvement or decrease in KCCQ is the minimum difference that is clinically relevant.<sup>7</sup>

5. Transvalvular mean gradient at 1 year (superiority). The hypothesis of interest is

$$H: \mu_{TAVR} < \mu_{SAVR}$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the average mean gradient at 1 year, measured in mmHg. This objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.975.

6. Effective orifice area at 1 year (superiority). The hypothesis of interest is

$$H: \mu_{TAVR} > \mu_{SAVR}$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the mean effective orifice area at 1 year, measured in  $\text{cm}^2$ . This objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.975.

7. Change in Kansas City Cardiomyopathy Questionnaire (KCCQ) score from baseline to 30 days (superiority). The hypothesis of interest is

$$H: \mu_{TAVR} > \mu_{SAVR}$$

where  $\mu_{TAVR}$  and  $\mu_{SAVR}$  denote the mean changes in the KCCQ score from baseline to 30 days. For subjects with KCCQ score at both baseline and 30 days, the change in KCCQ will be calculated as  $\text{KCCQ}_{30\text{day}} - \text{KCCQ}_{\text{baseline}}$ . The objective will be evaluated using a Bayesian version of a two-sample t-test. The posterior probability  $P(H | \text{data})$  will be calculated and compared to a threshold of 0.975.

## 6.10. Heterogeneity/Poolability

Incidence of the primary endpoint will be evaluated for poolability across key subgroups. In particular, the primary endpoint will be examined for differences in outcome between gender, and need for revascularization. Tests for these outcomes will be performed to evaluate potential interactions between treatment and gender, and between treatment and need for revascularization.

### 6.10.1. Geography Poolability Analysis

Poolability analyses may be conducted either with frequentist or Bayesian statistical methods. The analysis descriptions below are written with the language of frequentist methods. If Bayesian methods are used, analogous models will be employed, with non-informative prior distributions, and statements below such as “significant at the 0.15 level” can be understood to mean if the 85% equal-tailed Bayesian credible interval (BCI) for the parameter of interest excludes 0.

#### 6.10.1.1. Primary Endpoint by Geography

The interaction between geography (US vs. Other) and treatment on the probability of death or disabling stroke at 24 months will be compared using a Cox proportional hazard model. This analysis will be performed on the AT set. If the resulting test is significant at the 0.15 level, further exploratory analysis will attempt to identify covariates that may explain treatment effect differences between the regions. Otherwise, the data will be considered to be poolable across geographies.

#### 6.10.1.2. Univariate Covariate Analysis

If in the analysis of primary endpoint by geography, the resulting p-value is  $\leq 0.15$ , then the following baseline characteristics will be examined individually as potential predictors of death or disabling stroke:

- Gender

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**



- Age
- Baseline NYHA
- STS score
- Baseline LVEF
- Hypertension
- Diabetes
- Coronary artery disease
- Prior stroke
- Prior MI
- Prior PCI

### **6.10.1.3. Multivariable Analysis**

Along with geography and treatment-by-geography interaction, any covariates with a 0.20 significance level from the Univariate Covariate Analysis will be included in a Cox proportional hazard model. If this multivariable regression does not result in a significant (level 0.15) geography by treatment interaction after adjustment for these baseline factors, then outcome results will again be considered poolable across geographies. If geography by treatment interaction is still significant (level 0.15) after adjustment for these factors, results will be presented by geography and the clinical significance of these differences will be assessed.

### **6.10.2. Site Poolability Analysis**

If the geographic regions are considered to be poolable, then the Site Poolability Analysis will be performed on all sites, regardless of geography; the geographic regions (US vs. Other) will not be taken into account.

Should, however, the geographic regions US and Other not be considered to be poolable, then for both of these regions a separate site poolability analysis will take place.

#### **6.10.2.1. Pooling of Small Sites**

Sites should contribute to at least 5 treatment and at least 5 control subjects to the AT set. If this is not the case, the site is considered a “small site”; small sites will be ordered by the date of first procedure in the AT set. Starting with the first “small site”, a pseudo-site will be created by adding subjects from successive “small sites”. Once the number of subjects (treatment and control subject together) reaches or exceeds the size of the median enrollment (treatment and control subject together) of the “large sites”, then a second pseudo-site will be created, beginning with the next site not already included in the first pseudo-site. Additional pseudo-sites, if needed, would be created in the same manner.

If the geographic regions (US vs. Other) are considered to be poolable then the pseudo-sites can consist of small sites of both geographic regions. If however the regions are not considered to be poolable, then specific pseudo-sites for both regions need to be created. A separate site poolability analysis will be performed for both regions.

#### **6.10.2.2. Primary Endpoint by Site**

The interaction between site or pseudo-site and treatment on the probability of death or disabling stroke will be compared using a Cox proportional hazard model. This analysis will be performed in the AT set. If the resulting test is significant at the 0.15 level, further exploratory analysis will be conducted. One analysis will implement a random-effects model for the primary endpoint that includes site as a random effect. Another analysis will attempt to identify covariates that may explain treatment effect differences among the sites. Otherwise, the data will be considered to be poolable across study sites.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

### 6.10.2.3. Multivariable Analysis

If in the analysis for Primary Endpoint by Site, the resulting test is significant at the 0.15 level, then the covariates significant at the 0.20 level from Univariate Covariate Analysis (6.10.1.2) will be included along with site in a Cox proportional hazard model. If this multivariable regression does not result in a significant (level 0.15) site by treatment interaction after adjustment for these baseline factors, then outcome results will again be considered poolable across study sites. If site by treatment interaction is still significant (level 0.15) after adjustment for these factors, results will be presented by site and the clinical significance of these differences will be assessed.

---

## 7. Additional Analysis

There will be approximately 50 AT subjects from Japan in this study. Once the primary objective demonstrates non-inferiority (either at one of the interim analyses, or at the final analysis), an additional report specifically for these Japanese subjects will be prepared as requested by PMDA, which includes descriptive statistics for the primary endpoint and all secondary endpoints. In addition, poolability analysis (Japan vs. non-Japan data) will be performed.

---

## 8. Validation Requirements

Statistical programming for the analysis datasets, primary endpoint, 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.

---

## 9. References

<sup>1</sup> Reardon MJ, Adams DH, Kleiman NS, Yakubov SJ, Coselli JS, Deeb GM, Gleason TG, Lee JS, Hermiller JB, Jr., Chetcuti S, Heiser J, Merhi W, Zorn GL, 3rd, Tadros P, Robinson N, Petrossian G, Hughes GC, Harrison JK, Maini B, Mumtaz M, Conte JV, Resar JR, Aharonian V, Pfeffer T, Oh JK, Qiao H and Popma JJ. *2-Year Outcomes in Patients Undergoing Surgical or Self-Expanding Transcatheter Aortic Valve Replacement. Journal of the American College of Cardiology.* 2015;66:113-21.

<sup>2</sup> Farrington CP and Manning G. *Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. Stat Med.* 1990;9:1447-54.

<sup>3</sup> Hintze J. NCSS and PASS, *Number Cruncher Statistical Systems.* Kaysville, UT. [www.ncss.com](http://www.ncss.com). 2014.

<sup>4</sup> Jennison C and Turnbull BW *Group Sequential Methods with Applications to Clinical Trials:* Boca Raton: Chapman and Hall; 2000.

<sup>5</sup> Reardon MJ, et al., *Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients.* N Engl J Med 2017; March 17 doi: 10.1056/NEJMoa1700456.

<sup>6</sup> Wilber DJ, Pappone C, et al., *Comparison of Antiarrhythmic Drug Therapy and Radiofrequency Catheter Ablation in Patients with Paroxysmal Atrial Fibrillation: A Randomized Controlled Trial* JAMA. 2010;303(4):333-340.

<sup>7</sup> Spertus J, et al, *Monitoring clinical changes in patients with heart failure: A comparison of methods,* American Heart Journal 2005;150:710-715.

This document is electronically controlled. Printed copies are considered uncontrolled.

**Confidential**

Transcatheter Aortic Valve Replacement With the Medtronic Transcatheter Aortic Valve Replacement System In Patients at Low Risk for Surgical Aortic Valve Replacement Continued Access Trial

Clinical Trials.gov Identifier: NCT02701283

Statistical Analysis Plan

Document Date: 13 MAR 2019 (Version 1.0)

## Medtronic Statistical Analysis Plan

<b>Clinical Investigation Plan Title</b>	Transcatheter Aortic Valve Replacement (TAVR) in Patients with the Medtronic Transcatheter Aortic Valve Replacement System (TAVR) in Patients at Low Risk for Surgical Aortic Valve Replacement (SAVR) Continued Access Trial Addendum
<b>Clinical Investigation Plan Identifier</b>	10831957DOC
<b>Sponsor/Local Sponsor</b>	Medtronic, Inc Coronary and Structural Heart Clinical 8200 Coral Sea St NE, MVS 66 Mounds View, MN 55112 United States
<b>Document Version</b>	Version 1.0

### Confidentiality Statement

The information contained in this document is confidential and the proprietary property of Medtronic. Any distribution, copying, or disclosure without the prior written authorization of Medtronic is strictly prohibited. Persons to whom the information is disclosed must know that it is confidential and that it may not be further disclosed by them.



## Table of Contents

<b>1. Version History .....</b>	<b>3</b>
<b>2. List of Abbreviations and Definitions of Terms.....</b>	<b>3</b>
<b>3. Introduction.....</b>	<b>3</b>
<b>4. Study Objectives.....</b>	<b>4</b>
<b>5. Investigation Plan .....</b>	<b>4</b>
<b>6. Determination of Sample Size .....</b>	<b>5</b>
<b>7. Statistical Methods .....</b>	<b>5</b>
7.1. Disposition of Subjects.....	5
7.2. Analysis Sets.....	6
7.3. General Methodology.....	6
7.4. Handling of Missing, Unused, and Spurious Data and Dropouts.....	6
7.5. Demographic and Other Baseline Characteristics .....	7
7.6. Interim Analyses.....	7
7.7. Evaluation of Objectives .....	7
7.8. Changes to Planned Analysis.....	8
<b>8. Validation Requirements .....</b>	<b>9</b>
<b>9. References.....</b>	<b>9</b>
<b>10. Statistical Appendices.....</b>	<b>9</b>

## 1. Version History

Version	Summary of Changes	Author(s)/Title
1.0	<ul style="list-style-type: none"> <li>Not Applicable, New Document</li> </ul>	Jian Huang, Principal Statistician

## 2. List of Abbreviations and Definitions of Terms

Abbreviation	Definition
TAVR	Transcatheter Aortic Valve Replacement
SAVR	Surgical Aortic Valve Replacement
TAV	Transcatheter Aortic Valve
TVT-R	Transcatheter Valve Therapies Registry
CMS	Centers for Medicare & Medicaid Services
KCCQ	Kansas City Cardiomyopathy Questionnaire
RDC	Remote Data Capture
TTE	Transthoracic Echocardiography
ICF	Informed Consent Form
MDCT	Multi-Detector Computed Tomography
PCI	Percutaneous Coronary Intervention
DTL	Delegated Task List
eCRFs	Electronic Case Report Forms
SID Number	Subject Identification Number

## 3. Introduction

This Statistical Analysis Plan has been designed to document, before data are analyzed, the rationale for the study design, and the planned analyses that will be included in study reports. This statistical analysis plan is developed based on the Clinical Investigational Plan (CIP) Addendum.

The Medtronic TAVR in Low Risk Patients Trial is a multi-center, prospective, randomized, interventional pre-market trial. The continued access phase of the trial permits the continued use of the TAVR therapy in patients at low surgical risk after the conclusion of the randomized phase and before receiving FDA approval of the IDE.

The continued access trial is a single-arm, non-randomized study. The purpose of this trial is to further evaluate the safety and effectiveness of the Medtronic TAVR System in patients with severe aortic stenosis at low risk for SAVR. Data from the continued access study will be used to additionally support regulatory submissions for the TAVR in Low Risk Patients study in the United States and other regions where applicable.

## 4. Study Objectives

---

The primary objective of the trial is to evaluate the safety and effectiveness of the Medtronic TAVR system as measured by the rate of all-cause mortality or all stroke at 1 year in subjects who have a low predicted risk of operative mortality for SAVR. Annual clinical summaries will report long term follow-up and monitor adverse events (as collected through the TVT Registry (TVT-R) and Centers for Medicare & Medicaid Services (CMS) claims data) through 10 years post implant.

## 5. Investigation Plan

---

The continued access phase of the Low Risk trial is a multi-center, prospective, non-randomized, interventional trial conducted in the United States. The maximum sample size is not expected to exceed 3660 subjects with an attempted implant, this maximum is based on enrollment during the randomized phase of the trial and will be distributed in semi-annual allocations of 732 attempted implants per 6 months. No site will enroll more than 15% of the semi-annual allocation (i.e., no site will implant > 110 subjects during any given allocation period) without prior authorization from Medtronic. Subjects who exit from the trial after implantation will not be replaced.

All enrolled, qualified, and heart team approved low risk subjects will be implanted with the Medtronic Evolut R or Evolut PRO system. Subjects that undergo an attempted implant or are implanted will be followed via TVT-R through 1 year with assessments at pre-, peri- and post-procedure, discharge, 30 days and 1 year. Additional clinical data available from CMS data claims will be reported out to 10 years.

Subjects will be consented for follow-up through 10 years. The enrollment period for this continued access study will be through completion of FDA review for the PMA supplement for

---

the TAVR Low indication; therefore, the estimated total duration of the trial (first subject enrolled to last subject completing his/her last follow-up exam) is estimated to be 12 years.

## 6. Determination of Sample Size

---

The sample size of the continued access phase of the TAVR in Low Risk Patients Trial will be determined based on FDA approval. The sample size is not expected to exceed 3660 attempted implant subjects at up to 61 hospitals in the United States.

This is not a hypothesis-driven study, therefore the sample size for the analysis was not determined by statistical sample size methods.

## 7. Statistical Methods

---

### 7.1. Disposition of Subjects

This trial will utilize an Oracle RDC system for subject tracking that is the property of Medtronic. Electronic records in Oracle RDC will include subject consent, date of heart team approval, date of attempted implant or implant, the name of the implanters and information identifying the devices used. This trial will also utilize the TVT-R database system for data collection through one year follow up for all attempted implant and implanted subjects.

After all one year follow-up is complete, standard of care clinical data available through claims data submitted to CMS will be provided for 2- through 10-year follow-up. ACC/STS will link CMS claims data with subject records in TVT-R. The CMS and TVT-R linked data is provided to Medtronic with adverse event tables for 2- through 10-year follow-up reporting requirements.

Follow-up visits at 30-day and one year are conducted through TVT Registry. The distribution of follow up visits and visit compliance by one year will be summarized including number expected, visit completed, missed visit, deaths, withdrawal, lost to follow-up and visit pending.

When TVT Registry data is merged with CMS for longer-term follow-up only, subjects that are matched between the databases will be included in subject accountability and analysis. CMS data does not have specified follow-up visits and study exit data is not applicable. The number of subjects in the CMS and TVT-R linked data will provide the number of subjects at risk in the cohort and for each specified timepoint (1-year, 2 years, 3 years, through 10-years). The following will be summarized: expected, completed, deaths, and pending.



## 7.2. Analysis Sets

If the subject signs informed consent, meets all inclusion and none of the exclusion criteria, and the Heart Team determines the subject is suitable for the trial, the subject is assigned TAVR and then added to the approved population. Within the approved population the following analysis sets are distinguished:

- **The attempted implant set:** The attempted implant set consists of all Heart Team approved subjects with an attempted implant procedure, defined as when the subject is brought into the procedure room and any of the following have occurred: anesthesia administered, vascular line placed, TEE placed, or any monitoring line placed.
- **The implanted set:** The Implanted set consists of all the attempted implant subjects who are implanted with the Medtronic TAV device.

## 7.3. General Methodology

Descriptive statistics will be used to report study data. For continuous variables (e.g. age), the mean, median, standard deviation, minimum, maximum and interquartile ranges (IQR) will be presented. For categorical variables, the number and percentage of subjects in the category of interest will be presented.

For time to event variables, Kaplan-Meier analyses of event rates at 30 days, 1 year and annually through 10 years will be presented. For these analyses, the time points will correspond to 30 days, 365 days, 730 days, and annually through 10 years (365×3, 365×4, 365×5, etc.) days post implantation respectively. At each time point with data, the product limit estimate of the event, the number of subjects at risk, the number of subjects with events, and the 95% confidence interval will be presented. For subjects without an event, the date of censoring will be the latest date of all follow-up visits, assessments, and events (including death in those endpoints where death is not the endpoint).

## 7.4. Handling of Missing, Unused, and Spurious Data and Dropouts

Every effort will be undertaken to minimize missing data. In time-to-event outcomes drop-outs will be censored at the time of discontinuation, consistent with the Kaplan-Meier approach.

Data from subjects that cannot be analyzed for a specific variable will be displayed as missing in the relevant summary tables. In this manner, all data for a specific variable are accounted for. No statistical techniques will be used to impute missing data.

In the case of partial dates, if only the month and year are known, the event or assessment will be analyzed as if it occurred on the 15th of that month. If only the year is known, the event or assessment will be analyzed as if it occurred on June 30th of that year. These resolutions of partial dates are subject to the restrictions that events must occur no earlier than the procedure date.

## 7.5. Demographic and Other Baseline Characteristics

Baseline demographic and clinical variables will be summarized for the attempted implant set. These include, but are not limited to, the baseline demographics and patient characteristics such as age, gender, race, STS score; the medical history such as myocardial infarction, stroke, peripheral arterial disease, atrial fibrillation/flutter, porcelain aorta, hostile chest; and the baseline echo measurements such as aortic valve area, mean gradient, left ventricular ejection fraction, aortic regurgitation, and mitral regurgitation.

All continuous variables will be summarized as means, medians, standard deviations, interquartile ranges, minima and maxima, and categorical variables will be summarized as frequencies and percentages.

## 7.6. Interim Analyses

No interim analyses will be conducted for this observational study.

## 7.7. Evaluation of Objectives

### 7.7.1. Primary Endpoint

All-cause Mortality or all Stroke Rate at 1-year post procedure.

## 7.7.2. Secondary Safety Endpoints

- Composite of death, all stroke, life-threatening bleed, or major vascular complication at 30 days
- New permanent pacemaker implantation at 30 days
- Prosthetic valve endocarditis at one year
- Prosthetic valve thrombosis at one year
- All stroke at one year
- Life-threatening bleed at one year
- Valve-related dysfunction requiring repeat procedure at one year (surgical or interventional therapy)

## 7.7.3. Secondary Effectiveness Endpoints

- Quality of Life as assessed by Kansas City Cardiomyopathy (KCCQ) change from baseline at one year
- Device Success (intra-procedure)

The endpoints are measured in accordance with the TVT-R definitions. The analysis will be descriptive, and no statistical hypothesis test will be performed.

Kaplan-Meier estimates at 30 days and 12 months will be provided for primary and secondary safety endpoints. In addition, Kaplan-Meier estimates will be provided for all-cause mortality or all stroke at 2 years and annually up to 10 years. Device success will be summarized using frequencies and percentages for the implanted set. KCCQ at baseline, 30 days, and 12 months, and KCCQ change from baseline at 30 days and 12 months will be summarized as means, medians, standard deviations, interquartile ranges, minima, maxima and 95% confidence intervals. The attempted implant set will serve as the primary analysis cohort for all the objectives (except for device success). Device success analysis will be based on the implanted set.

## 7.8. Changes to Planned Analysis

The planned analyses in this SAP are aligned with the planned analyses noted in the CIP addendum.

## 8. Validation Requirements

---

Level 1 validation (independent validation) will be used for the analysis datasets and the all endpoints.

## 9. References

---

NA

## 10. Statistical Appendices

---

There are no statistical appendices for this study.